(1) Publication number:

0 293 274 A1

1

EUROPEAN PATENT APPLICATION

2) Application number: 88400790.7

2 Date of filing: 31.03.88

(a) Int. Ci.4: C 12 N 15/00

A 61 K 39/29

- 30 Priority: 31.03.87 JP 78313/87 04.06.87 JP 140586/87 10.11.87 JP 283990/87
- Date of publication of application: 30.11.88 Bulletin 88/48
- Designated Contracting States:
 CH DE FR GB IT LI NL SE
- Applicant: MITSUBISHI CHEMICAL INDUSTRIES
 LIMITED
 5-2, Marunouchi 2-chome Chiyoda-ku
 Tokyo 100 (JP)
- inventor: Takahashi, Kazuhiro 303 Corpo Omodaka 29 ichijoji-Somedono-cho Sakyo-ku Kyoto City Kyoto Prefecture (JP)

Shibui, Tatsurou M606-11 San White 6-15-4 Minami-Naruse Machida City Tokyo (JP) Kamizono, Michiru F202-24 San White 2-8-11 Tauruma Machida City Tokyo (JP)

Matsui, Rie A211-23 San White 1-26-31 Ogawa Machida City Tokyo (JP)

Teranishi, Yutaka 1-14-2 Takane Sagamihara City Kanagawa Prefecture (JP)

Nakanishi, Shigetada 517-116 Iwakura-Nagatani-cho Sakyo-ku Kyoto City Kyoto Prefecture (JP)

Kitamura, Naomi 2-272 Yakumo-Higashi-machi Moriguchi City Osaka Prefecture (JP)

Representative: Gutmann, Ernest et al S.C. Ernest Gutmann - Yves Plasseraud 67, boulevard Haussmann F-75008 Paris (FR)

The title of the invention has been amended (Guidelines for Examination in the EPO, A-III, 7.3).

The application is published incomplete as filed (Article 93 (2) EPC). The point in the description or the claim(s) at which the omission obviously occurs has been left blank.

- Machine the second of the seco
- SP Provided herein are a DNA fragment which contains a base sequence coding for a non-A non-B hepatitis-specific antigenic protein occurring in cells of the liver affected with non-A non-B hepatitis, an expression vector in which said DNA fragment is inserted into a cloning site present downstream from a promoter thereof, a transformant obtained by introducing said expression vector into a host, and a process for producing said antigenic protein which comprises providing said expression vector, transforming a host with said expression vector, culturing the transformed host and collecing the protein produced therein.

EP 0 293 274 A1

Bundesdruckerei Berlin

Description

5

10

25

30

DNA FRAGMENTS CODING FOR ANTIGENS SPECIFIC TO NON-A NON-B HEPATITIS, EXPRESSION VECTORS CONTAINING SAID DNA FRAGMENTS, TRANSFORMANTS AND PROCESS FOR PRODUCING SAID ANTIGENS

BACKGROUND OF THE INVENTION

Field of the Invention:

The present invention generally relates to the production of an antigen specific to non-A non-B hepatitis by recombinant DNA technology. More particularly, it relates to a DNA fragment coding for an antigen specifically occurring in a host affected with non-A non-B hepatitis, an expression vector containing such a DNA fragment, a host transformed with such an expression vector, as well as a process for producing said antigen specific to non-A non-B hepatitis by culturing such a transformant.

15 Description of the Prior Art:

Among viral hepatitises, the viral entitles of hepatitis type A and type B have been found and, accordingly, it has now become possible to diagnose such diseases by immunological methods.

Still another type of hepatitis different from the types A and B, which is called non-A non- β type hepatitis, is said to be over 90% of post-transfusion hepatitis: refer to NIPPON RINSHO (Japan Clinic), 35, 2724 (1977); J. Blol. Med., 49, 243 (1976). The pathogenic virus of the non-A non-B type hepatitis, however, has not yet been identified. Only one fact which has already been established is potential infection of human hepatitis type non-A non-B virus to chimpanzee: refer to Lancet I, 459 (1978); ibid., 463 (1978).

Many workers have done various investigations for searching an antigen-antibody system related to the non-A non-B hepatitis by using mainly sera from patients affected with the disease; nevertheless, no definite system has been found. Under these circumstances, the diagnosis of non-A non-B hepatitis should inevitably be effected by so-called exclusion diagnosis: that is, whether or not the hepatitis of a patient is type A or type B or other hepatitis due to a virus known to cause hepatopathy, for example, CMV, HSV, EBV, etc., Is first determined; and if not, the patient's hepatitis is diagnosed as non-A non-B type. Thus, such a diagnosis of non-A non-B hepatitis will require much time and labor.

An antigenic protein specific to non-A non-B hepatitis and useful for the direct diagnoses of the hepatitis has been purified from human and chimpanzee hepatocytes affected with non-A non-B hepatitis, and a monoclonal antibody specific to the antigen and useful for the treatment of the non-A non-B hepatitis has also been proposed; refer to Japanese Patent Application Laying-open (KOKAI) Nos. 176856/86 and 56196/86.

A large amount of such an antigenic protein specific to non-A non-B hepatitis should be required when such a protein is to be employed, for example, as a diagnostic agent. However, it is not always appropriate to purify such a large amount of the antigenic protein from chimpanzee hepatocytes affected with non-A non-B hepatitis.

On the other hand, in order to detect a gene coding for a specific antigen of non-A non-B hepatitis by nucleic acid hybridization and, further, to produce such an antigen specific to non-A non-B hepatitis by the recombinant DNA technology, it is essential to obtain a gene fragment coding for the antigenic protein specific to the non-A non-B hepatitis.

SUMMARY OF THE INVENTION

The present inventors have made great efforts to produce such a specific antigenic protein in a large amount by genetic engineering techniques, and finally isolated a gene fragment coding for the antigenic protein specific to non-A non-B hepatitis, said gene fragment being useful for the production of such antigens. Further, the inventors have successfully constructed an expression vector containing said gene fragment. Thus, the present invention has now been attained.

It is an object of the invention to provide a DNA fragment which contains a base sequence coding for an antigen specifically occurring in a host cell affected with non-A non-B hepatitis or an antigenic protein specific to non-A non-B hepatitis having physiological activities equivalent to those of said specifically occurring antigen.

Another object of the invention is to provide an expression vector having said DNA fragment introduced thereinto at a cloning site downstream from a promoter of the vector.

A still another object of the invention is to provide a transformant obtained by transforming a host cell with said expression vector.

A further object of the invention is to provide a process for producing such an antigen specific to non-A non-B hepatitis by culturing said transformant.

BRIEF DESCRIPTION OF THE DRAWINGS

Other objects and advantages of the present invention will be apparent from the following detailed description with reference to the attached drawings, in which:

Figs. 1a-1e show the base sequence coding for an antigenic protein specific to non-A non-B hepatitis;

Fig. 2 shows the base sequence of a hybrid promoter Pac;

Figs. 3a-3c show the base sequence of a cDNA fragment obtained in Example 1 described hereinbelow, together with deduced amino acid sequence;

Figs. 4a-4c show the base sequence of cDNA containing the full length gene sequence of an antigenic protein specific to non-A non-B hepatitis, which cDNA was obtained in Example 2 described hereinbelow, the base sequence 57-1388 thereof coding for the antigenic protein specific to non-A non-B hepatitis;

Fig. 5 schematically illustrates the construction of a plasmid pCV44H;

Fig. 6 schematically illustrates the construction of a plasmid pCV44B; and

Fig. 7 schematically illustrates the construction of a plasmid pCZ44.

DETAILED DESCRIPTION OF THE INVENTION

The present invention will be described in detail hereinbelow.

According to one aspect of the invention, a DNA fragment is provided which contains a base sequence coding for an antigenic protein occurring specifically in hepatocytes affected with non-A non-B hepatitis.

Such a DNA fragment of the invention may be prepared in the following manner.

First, a liver tissue specimen derived from a human or chimpanzee individual affected with non-A non-B hepatitis is homogenized in an aqueous solution of guanidinium thiocyanate and then subjected to cesium chioride equilibrium dusity gradient centrifugation according to Chirgwin et al. method (Biochemistry, 18, 5294-5299 (1979)) to separate total RNA as a precipitate. After separation, the total RNA is purified by phenol extraction and ethanol precipitation.

"Individuals affected with non-A non-B hepatitis" used as sources of liver tissue specimens in the invention may include those affected with so-called type D hepatitls, which has recently been named.

It is known that mRNA of an antigen gene generally has a poly-A chain. Thus, the total RNA is subjected to oligo(dT) cellulose column chromatography in a conventional manner and poly(A)-containing RNA (poly A+ RNA) is isolated as mRNA material.

A cDNA library corresponding to the poly A+ RNA is then obtained from the mRNA material according to the random primer method (Y. Ebina et al., Cell, 40, 747-758 (1980)): Thus, a number of DNAs complementary to the mRNA material are randomly synthesized using any primer of e.g. about 6 bases and a reverse transcriptase.

The cDNA is methylated with a DNA methylase, e.g. EcoRI methylase, to protect cleavage sites present in the cDNA capable of being cleaved by a corresponding restriction enzyme, e.g. EcoRl. A DNA linker containing the corresponding restriction enzyme cleavage sites at both ends, e.g. EcoRI linker (CGAATTCG), is added to the methylated cDNA and, then, this cDNA is digested with the restriction enzyme, e.g. EcoRI.

The digested cDNA is then cloned into a cloning vector such as a plasmid or a λ phage. For example, the cDNA may be introduced into EcoRI site of \(\text{\gamma} \text{t1 DNA, which is an expression cloning vector: refer to R.A.} \) Young et al., Pro. Nati. Acad. Sci. U.S.A., 80, 1194-1198 (1983). The cDNA will be inserted into the β-gal gene on the Agt 11 phage. Thus, expression of the cDNA can be easily verified by the production of a fused protein with β-galactosidase due to induction of the expression by the lactose operon promoter of said phage when E. coli transfected with said phage is cultured in a medium containing IPTG (isopropylthlo-β-D-galactopyranoside).

The λgt 11 phage incorporating the cDNA is then introduced into E. coil by Tomizawa et al. method in Experimental Procedures for Bacteriophages", pp. 99-174, published May 30, 1970 by Iwanami Shoten (Japan). The thus transfected microorganism is cultured in an IPTG-containing medium.

The thus formed plaques can be easily selected by an immunological screening method using a monoclonal antibody specifically directed to non-A non-B hepatitis to obtain a desired cDNA. Such a monoclonal antibody which can be used in the immunological screening method may be prepared according to the methods described in Japanese Patent Application Laying-open Nos. 176856/86 and 56196/86. The screening methods used may include the western blotting technique described in these applications.

The plaques positive in the immunological screening test are selected to proliferate the phage by Tomizawa et al. method. DNA is purified from the grown phage by T. Maniatis et al. method in "Molecular Cloning", Cold Spring Harbor Laboratory, pp. 85 et seq. (1982), and cleaved with a suitable restriction enzyme such as EcoRI. The thus purified and digested DNA fragments can be used to determine the base sequence of a desired cDNA segment according to Maxam and Gilbert method in Methods in Enzymology, 65, 499-560 (1980); or alternatively, after further cloning the DNA fragments into M13 phage, the base sequence of such a desired cDNA segment can be determined according to the dideoxy method: Sanger et al., Proc. Natl. Acad. Sci. U.S.A., 74, 5463 (1977).

Thus, a cDNA fragment coding for an antigen specific to non-A non-B hepatitis can be obtained. However, such a DNA fragment may usually be only a portion of the gene coding for the non-A non-B hepatitis-specific antigen.

A full length cDNA coding for such a non-A non-B hepatitis-specific antigen may be obtained in the following

Poly A*-mRNA is isolated and purified in a manner similar to that described above. From the poly A*-mRNA a cDNA library is obtained according to Okayama-Berg vector-primer method: Molecular and Cellular Biology, 2, 161-170 (1982).

A plasmid containing such a cDNA thus prepared is used to transform E. coil by any conventional method, for instance, the method D. Hanahan: J. Mol. Biol., 166, 557 (1983). The transformant ampicillin-resistant

3

10

5

15

20

25

*3*0

35

45

50

55

∞

65

strains are collected and screened by the colony hybridization method using the aforementioned DNA fragment as a probe. Such a probe may preferably be prepared by either the strepto-avidin method, or the nick translation method using photobiotinnucleic acids and 32P-nucleic acids.

The thus selected colonies containing a cDNA clone are cultured. Plasmid DNA is obtained from the cultured colony according to Birmboim et al. method (Nucleic Acid Res., 7, 1513 (1979)) and digested with a suitable restriction enzyme. The base sequence of a desired full-length cDNA segment is then determined according to the aforementloned Maxam and Gilbert method or, alternatively, after further cloning the digested DNA into M13 phage or pVC12 plasmid, such a base sequence is determined according to the above described Sanger et al. dideoxy method.

The base sequence of the full length DNA coding for an antigen specific to non-A non-B hepatitis is shown in Fig. 1, in which the symbol "-" just under the base sequence represents a corresponding base complementary to the respective base described just above each of the symbols.

Of course, DNA fragments which can be employed in the invention do not necessarily contain the same base sequence as shown in Fig. 1, but those DNA fragments in which a part of said base sequence shown in Fig. 1 has been substituted by at least one different base or deleted therefrom and those DNA fragments in which one or more additional bases have been added to the base sequence of Fig. 1 may also be included herein provided that such different DNA fragments may code for substances having physiological activities equivalent to those of the non-A non-B hepatitis-specific antigens encoded by the base sequence of Fig. 1.

According to another aspect of the invention, an expression vector is provided in which the aforementioned DNA fragment of the invention is inserted into a cloning site downstream from a promoter of this vector.

The expression vector of the invention contains a promoter in a position capable of controlling the transcription of a DNA fragment coding for a non-A non-B hepatitis-specific antigen obtained by the aforementioned method. The promoters used in the invention may be any promoter capable of expressing the DNA fragment in a host, and preferably of controlling the transcription of the fragment.

When a host used is a microorganism such as Escherichia coli, Bacillus subtilis, etc., the expression vector of the Invention may preferably comprise a promoter, a ribosome binding sequence, a gene for a non-A non-B hepatitis-specific antigen, a transcription termination factor, and a gene controlling the promoter.

The promoter used may include those derived from E. coli, phage, etc., for example, tryptophan synthase operon (trp), lactose operon (lac), lipoprotein (lpp), recA, lambda phage PL, PR, T5 early gene P25, P26 promoter, which may also be prepared by chemical synthesis. Also included herein are hybrid promoters such as tac (trp:lac), trc (trp:lac) and Pac (phage:E. coli) shown in Fig. 2.

The ribosome binding sequence may be derived from E. coli, phage, etc., but preferably may be those synthetically prepared, for example, those containing a consensus sequence such as

AGGAGGTTTAA.

SD sequence

The gene for a non-A non-B hepatitis-specific antigen may be directly employed without any modification. Preferably, an unnecessary base sequence (non-coding region) may be deleted by site-directed mutagenesis: BIO TECHNOLOGY, July, 636-639 (1984).

A transcription termination factor may not always be required in the expression vector of the invention. Preferably, the instant vector may contain a p-independent termi nator, for example, ipp terminator, trp operon terminator, ribosomai RNA gene terminator, etc.

The expression vector may be derived from any conventional plasmid. Preferably, it may be derived from such a plasmid as replicating itself in E. coll or Bacillus subtilis, for example, pBR322- or pUB110-derived plasmid.

Desirably, these factors required for expression are arranged in the expression plasmid in the order of the promoter, the SD sequence, the structural gene of a non-A non-B hepatitis-specific antigen, and the transcription termination factor from 5' to 3'. A repressor gene required to control the transcription, a marker gene such as drug-resistant gene, and a plasmid replication origin may be arranged in any order in the expression vector.

The expression vector of the invention may be introduced into a host by any conventional method for transformation of E. coli, e.g., one described in Molecular Cloning, 250-253 (1982), or of Bacillus subtills, e.g., one described in Molec. Gen. Genet., 168, 111-115 (1979) or Proc. Nat. Acad. Sci. U.S.A., 44, 1072-1078 (1958).

The resulting transformant may be cultured in any conventional medium, e.g. one described in Molecular Cloning, 68-73, (1972), at a temperature in the range of 28 to 42°C in both cases of E. coil and Bacillus subtilis. Preferably, it may be cultured at a temperature in the range of 28 to 30°C where no expression of heat shock proteins may be induced.

The desired protein thus produced may be easily purified from the host in conventional procedures. For example, the host cell may be crushed by lysozyme-surfactant or ultra-sonication, and the insoluble fractions which contain the desired non-A non-B hepatitis-specific antigen may be then collected by centrifugation,

4

35

10

20

25

45

50

solubilized in a surfactant such as 0.01% SDS, and subjected to column chromatography using a monoclonal antibody (Japanese Patent Application Laying-open (KOKAI) Nos. 56196/86 and 176856/86.

When an eukaryotic cell such as an animal cell is employed as a host, the expression vector of the invention is preferably as follows:

5

10

15

20

25

*3*0

35

45

50

55

60

65

The promoters used in the vector of the invention for the expression in eukaryotic cells may herein include SV40 early and late promoters; promoters of apolipoprotein E and A-I genes; promoter of heat shock protein gene (Proc. Natl. Acad. Sci. U.S.A., 78, 7038-7042 (1981)); promoter of metallothionein gene (Proc. Natl. Acad. Sci. U.S.A., 77, 6511-6515 (1980)); HSV TK promoter; adenovirus promoter, such as Ad2 major late promoter (Ad2 MLP); LTR (long terminal repeat) of retrovirus; etc. SV40 promoter and promoter of metallothionein gene are preferred.

The expression vector of the invention may contain a splice sequence comprising 5' splice junction donor site, an intron and 3' splice junction acceptor site. A common base sequence is found at all the splice junction sites (exonintron junction sites); so-called GT/AG rule that any intron region always starts from two bases GT at the donor site and terminates at two bases AG of the acceptor site has been established.

The expression vector of the Invention may contain one or more splice sequences as mentioned just above. Such splice sequences may be positioned upstream or downstream of the structural gene for a non-A non-B hepatitis-specific antigen.

lilustrative examples of such splice sequences may include those DNA sequences found in exons 2 and 3 of rabbit β -globin gene (Science, $\underline{26}$, 339 (1979)) and mouse methallothionein-I gene containing the promoter, exons 1, 2 and 3 and introns A and B of methallothionein gene (Proc. Natl. Acad. Sci. U.S.A., 77, 6513 (1980)). The 5' and 3' splice sites may be derived from the same or different gene; for example, a sequence in which 5' splice site contained in adenovirus DNA is linked to 3' splice site derived from the gene of Ig variable region can be employed.

The expression vector of the invention also contains a polyadenylation site downstream from the structural gene of a non-A non-B hepatitis-specific antigen. Illustrative examples of the polyadenylation sites may include those derived from SV40 DNA, β -globin gene or methallothioneln gene. A combined site of the polyadenylation sites of β -globin gene and SV40 DNA may be employed in the Invention.

The expression vector of the invention may also contain a dominant selective marker permitting the selection of transformants. Selective markers which can be used herein may include DHFR gene imparting MTX (methotrexate) resistance to a host; tk gene of herpes simplex virus (HSV) which permits selection of tk strains transformed therewith in HAT medium; the gene for aminoglycoside 3'-phosphotransferase from E. coll transposon Tn5, which imparts to a host the resistance against 3'-deoxystreptamine antibiotic G418; bovine papilloma virus gene permitting morphological discrimination by piled up growth; and aprt gene.

Alternatively, animal cells transformed with the expression vector of the Invention may be selected by the cotransformation even though no selective marker is present in the vector. For this purpose, an animal cell is cotransformed with both the expression vector and a plasmid or other DNA containing a gene for such a selective marker and selected by a phenotypic trait of the gene.

Advantageously, the expression vectors may also contain a plasmid fragment having an origin of replication derived from a bacterium such as <u>E. coli</u>, since such vectors can be cloned in bacteria. Such plasmids may include pBR322, pBR327, pML, etc.

Illustrative examples of plasmid vectors used as sources of the expression vectors according to the invention may include pKCR (Proc. Nati. Acad. Sci. U.S.A., 78, 1528 (1981)), which contains SV40 early promoter, the splice sequence and polyadenylation site derived from rabbit β-globin gene, the polyadenylation site from SV40 early region, and the origin of replication and ampicillin resistant gene from pBR322; pKCR H2 (Nature, 307, 605 (1984)), in which the pBR322 portion of pKCR has been substituted by pBR327 fragment and the EcoRI site present in the exon 3 of rabbit β-globin gene has been converted into Hindlil site; and pBPV MT1 containing BPV gene and methallothioneln gene (Proc. Natl. Acad. Sci. U.S.A., 80, 398 (1983)).

Animal cells transformed with the expression vector of the invention may include CHO cells, COS cells, and mouse L cells, C127 cells and FM3A cells.

The introduction of the expression vector of the invention into an animal cell may be carried out by transfection, microinjection, etc. Most often, the transfection may employ CaPO₄: Virology, <u>52</u>, 456-467 (1973).

Animal cells transformed by introducing the expression vector of the invention may be cultured in a suspension or solid medium by conventional methods. The culture medium used is most often MEM, RPMI1640, etc.

Proteins produced in the transformed animal cells can be separated and purified in the almost same manner as in the case of microorganisms aforementioned.

As stated, the invention provides a transformant cell obtained by introducing the expression vector of the invention into a host cell.

Also provided according to the invention is a process for producing a non-A non-B hepatitis-specific antigen comprising culturing said transformant and collecting the produced and accumulated antigen.

As stated previously, a large amount of an antigenic protein specific to non-A non-B hepatitis is required when such a protein is to be utilized as a direct diagnostic agent. According to the present invention, such an antigenic protein can be produced with a low cost and a large scale without use of infected chimpanzee hepatocytes. Prior to the present invention, it has been difficult obtain such a large amount of a non-A non-B hepatitis-specific antigenic protein from hepatocytes of chimpanzees affected with non-A non-B hepatitis.

Further, the DNA fragment coding for an antigenic protein of non-A non-B hepatitis virus according to the present invention will be useful as a probe for detecting the gene of said antigenic protein by nucleic acid hybridization.

EXAMPLES

5

10

20

25

45

65

The following examples will be given by way of Illustration but these examples in no way limit the scope of the invention without departing the concept thereof.

EXAMPLE 1: Preparation of cDNA Fragment Coding for Antigenic Protein Specific to Non-A Non-B Hepatitis Poly(A)-containing RNA was prepared from chimpanzee liver according to the guanidine thiocyanate-lithium chloride method: Cathala et al., DNA, 2, 329 (1983).

The infected liver (5 g) was taken out from a chimpanzee affected with non-A non-B hepatitis and immediately frozen by liquid nitrogen. The frozen liver was added into a Waring blender together with liquid nitrogen and ground at 3,000 rpm for 2 minutes. The ground liver specimen was further ground by a Teflon homogenizer at 5 rpm in 100 ml of a solution: 5 M guanidine thiocyanate, 10 mM EDTA, 50 mM Tris-HCl (pH 7), 8% (v/v) β-mercaptoethanol. The thus solubilized material (20 ml) was slowly placed on 5.7 M CsCl solution (10 ml) contained in a centrifuge tube and centrifuged at 27,000 rpm for 20 hours in Hitachi RPS 28-2 rotor. The thus precipitated RNA was collected and dissolved in 10 ml of a solution: 0.1% sodium laurylsulfate, 1 mM EDTA, 10 mM Tris-HCl (pH 7.5). The RNA was extracted with phenol-chloroform and re covered by ethanol precipitation.

The thus obtained RNA (about 3.95 mg) was dissolved in 1 ml of a solution: 10 mM Tris-HCl (pH 8.0), 1 mM EDTA. The solution was incubated at 65°C for 5 minutes, and 5 M NaCl (0.1 ml) was added. The resulting mixture was subjected to chromatography on an oligo(dT) cellulose column (column volume of 0.5 ml, P-L Biochemical). The thus adsorbed poly(A)-containing mRNA was eluted with a solution: 10 mM Tris-HCl (pH

7.5), 1 mM EDTA. There was obtained about 100 µg of poly(A)-containing mRNA.

The thus obtained poly(A)+ mRNA (10 μg) was dissolved in 50 μl of RT buffer: 20 mM Tris-HCl (pH 8.8), 0.1 M KCl, 12 mM MgCl₂, 2 mM MnCl₂. To this solution, there was added 8 µg of random primer d(N)₆ (P-L Biochemical). The resulting mixture was heated at 95°C for 3 minutes to denature the materials, which was then cooled gradually to room temperature to anneal the random primer with the mRNA. To the annealed mixture, there were aded 10 mM 4NTP (10 µl) and reverse transcriptase (225 units) from TAKARA SHUZO (Japan), and then water was added so as to make the total volume of the mixture to 100 µl. Reaction was allowed to proceed at 42°C for one hour.

To the reaction mixture (50 µl), there were added 10 mM NAD (2 µl), 10 mM 4dNTP (10 µl), RNase H (5 units), E. coll ligase (1 unit), E. coll DNA polymerase I (6.3 units), and 10 x T4 DNA ligase buffer (10 μi; 0.1 M Tris-HCl, рН 7.5, 0.1 M DTT, 60 mM MgCl₂) to make the total volume to 100 µl. The mixture was allowed to react at 37°C

for one hour to synthesize a double stranded DNA.

The thus obtained double stranded DNA was extracted with an equal volume of water-saturated phenol. Phenol in the aqueous layer was removed with the aid of ether followed by ethanol precipitation. The precipitate thus obtained was dissolved in 50 μ l of water, and 10 \times T4 DNA polymerase buffer (10 μ l; 0.33 M Tris-acetic acid, pH 7.9, 0.66 M potassium acetate, 0.1 M magnesium acetate, 5 mM DTT), 10 mM 4dNTP (10 μί), and T4 DNA polymerase (6 units) were added to make the total volume to 100 μί. The mixture was reacted at 37°C for one hour. There was obtained a double stranded DNA having blunt ends, which was then extracted with phenol to remove proteins and purified by ethanol precipitation as described above. The thus purified DNA was then air dried.

To the purified DNA, there were added 50 mM Tris-HCI (pH 7.5), 1 mM Na₂EDTA, 5 mM DTT (20 μ), 100 μM S-adenosyl-L-methlonine (2 μl), and 1.8 mg/ml EcoRl methylase (0.2 μl). Reaction was effected at 37°C for 15 minutes, whereby methylating the EcoRi restriction enzyme cleavage site on the DNA fragment. The reaction

mixture was then heated at 70°C for 15 minutes to deactivate the enzyme.

To the reaction mixture, there was added 3'-phosphorylated EcoRi linker (GGAATTCC) in an amount of 100 molecules thereof per molecule of the synthetic DNA. There were further added 10 x T4 DNA ligase buffer (5 µl; 0.5 M Tris-HCl, pH 7.5, 60 mM MgCl₂, 10 mM DTT), 0.1 M ATP (5 μl), and T4 DNA ligase (5 units) to make the total volume to 50 μl. The resulting reaction mbxture was reacted at 4°C for 16 hours followed by heating at 70°C for 10 minutes to deactivate the enzyme. Then, $10 \times$ EcoRi buffer (10 μl; 15 M Tris-HCl, pH 7.5, 0.5 M NaCl, 60 mM MgCl₂), and EcoRl (100 units) were added to make the total volume to 100 µl, and the reaction mixture was reacted at 37°C for 2 hours to cut the linker. The reaction mixture was passed through Bio Gel A-50 (0.2 cm \times 32 cm, Bio RAD). Elution was effected by a buffer: 10 mM Tris-HCl (pH, 7.5), 6 mM MgCl₂. Excess EcoRl linker was removed and, thus, a double stranded cDNA having EcoRl sites at both ends thereof was purified.

To the thus obtained double stranded cDNA fragment having EcoRI sites at both ends, there were added gt 11 DNA (10 μ g) cleaved with EcoRI, 10 \times T4 DNA ligase buffer (10 μ l) as described above, 0.1 M ATP (10 μ l), and T4 DNA ligase (10 units) to make the total volume to 100 µl. The mixture was reacted at 4°C for 16 hours.

Thus, said double stranded cDNA fragment was inserted into λgt 11 DNA. The λ phage packaging kit (PROMEGA, Biotech) was used to Introduce said DNA into λ phage particle. The procedures for packaging were effected according to the instructions of the kit.

The λgt 11 phage having said DNA packaged thereinto was used to transfect E. coli strain Y1090 to form

plaques according to the conventional Tomizawa et al. methods described in *Experimental Procedures for Bacteriophages*, pp. 99-174, published May 30, 1970 by Iwanami Shoten (Japan). Among about 200,000 plaques, one positive clone was selected by Immunological screening as described hereinbelow. A monoclonal antibody used in the immunological screening was prepared by the method described in Japanese Patent Application Laying-open (KOKAI) No. 176856/86.

5

10

15

20

25

30

35

45

50

55

60

E. coli Y1090 (R.A. Young et al., Pro. Natl. Acad. Sci. U.S.A., 80, 1194-1198 (1983), which had been transfected with \(\lambda\)gt 11, was inoculated in a petri dish together with soft agar held at 42°C. The transfected cell was allowed to stand at 42°C for 5 hours. A nitrocellulose filter (S & S, BA-83, pore size of 0.2 μm) containing 10 mM IPTG was placed on the cell in the dish and incubation was effected at 37°C for 3-4 hours. This nitrocellulose filter was lightly rinsed with TBS buffer (10 mM Tris-HCl, pH 7.5, 50 mM NaCl), immersed in the TBS buffer (400 ml) containing 3% gelatine and shaked at 40°C for one hour. Thus, the nitrocellulose filter was blocked. Then, a monoclonal antibody (OD280 = 4.3) directed to a non-A non-B hepatitis-specific antigen was added to TBS buffer containing 1% gelatine with a dilution of 1/400. This mixture was put into a viryl bag together with the filter in a proportion of 2 ml of the mixture per filter, and reaction was allowed to proceed at room temperature for 16 hours. The reaction mixture was three times washed with TBS buffer (400 ml) containing 0.05% Tween 20 over 10 minutes. A labelled secondary antibody, anti-mouse IgG-PAP (horseradish peroxidase, Bio Rad) was added to TBS buffer containing 1% gelatine with a dilution of 1/1,000. This mixture and the filter were put into a vinyl bag with a proportion of 2 ml of the mixture per filter. Reaction was allowed to proceed at room temperature for 2 hours. The reaction mixture was three times washed with TBS buffer (400 ml) containing 0.05% Tween 20 over 10 minutes, in the same manner as described above. Color development was effected by dipping the filter and 4-chloro-1-naphthol (12 mg, Bio Rad) into 20 ml of TBS buffer containing hydrogen peroxide. After completion of the color development, the filter was thoroughly washed with water and put into a vinyl bag containing water. The bag was stored in a dark and cold place.

Thus, one positive plaque was obtained. The plaque was three times subjected to single plaque isolation. In each time, immunological screening was effected in the same manner as described above, verifying that the plaque was in fact positive.

The phage was then cultured in a large scale to purify the DNA in the following manner: First, E. coli Y1090 was cultured overnight in 10 ml of NZ medium prepared by adding NZ amine (10 g), NaCl (5 g) and 5 mM MgCl₂ to one liter of water followed by adjusting the pH to 7.2. The culture (1 ml) was transfected with the phage, with the m.o.i. (multiplicity of infection) being 0.1. The transfected culture was allowed to stand at 37°C for 10 minutes and then transferred to one liter of NZ medium. Shaking culture was effected at 37°C for 7-8 hours until the cells were lysed. Chloroform (5 ml) was added to the culture and shaking was continued for additional 30 minutes. The culture was subjected to centrifugation at 6,500 rpm for 10 minutes to remove cell debris.

NaCl (29 g) and polyethylene glycol (70 g) were added to and thoroughly dissolved in the obtained supernatant, and the solution was allowed to stand at 4°C overnight. The precipitate was collected by centrifugation at 6,500 rpm for 20 minutes, drained thoroughly, and dissolved in 20 ml of TM buffer: 10 mM Tris-HCl (pH 7.5), 5 mM MgCl₂ DNase I and RNase A were added to the solution, both with a concentration of 10 µg/ml, and the reaction was effected at 37°C for one hour. Chloroform (20 ml) was then added to the reaction mixture and stirred; thus, polyethylene glycol was distributed in the chloroform layer which was then separated from the aqueous layer. This aqueous layer was ultra- centrifuged at 28,000 rpm for 60 minutes. Thus, a pellet of phage particles was obtained.

This pellet was dissolved in TM buffer (1 ml) and subjected to CsCl density gradient centrifugation at 33.000 rpm for 20 hours. The resultant fraction containing the phage particles (p=1.45-1.50) was dialyzed overnight against TM buffer. Proteinase K was added to the dialyzate in an amount of 100 µg/ml and reaction was effected at 37°C for one hour. Thereafter, an equal volume of water-saturated phenol was added and phenol-extraction was gently effected. After centrifugation at 6,500 rpm for 10 minutes, the aqueous layer was removed, put into a dialysis tube, and dialyzed overnight against water at 4°C. Thus, about 5 mg of DNA was obtained.

Cleavage reaction of this DNA (100 µg) with EcoRI (100 units) in the aforementioned buffer (100 µl) at 37°C revealed that two cDNA segments of 390 bp and 345 bp were inserted into the phage DNA.

These two EcoRI fragments were re-cloned into EcoRI site of a cloning vector pUC 119. Base sequences of these DNA fragments were determined by the dideoxy method using commercially available primers CAGGAAACAGCTATGAC and AGTCACGACGTTGTA, respectively. The base sequence of the linking portion between these two DNA fragments was similarly determined by cutting this cDNA fragment at BamHI and EcoRV sites present therein with corresponding specific restriction enzymes, inserting the resulting BamHI-EcoRV DNA fragment between BamHI and Small sites of the plasmid pUC 119, and sequencing the fragment by the dideoxy method.

The base sequence of said cDNA fragment is shown in Fig. 3. This was a partial cDNA fragment of a gene coding for an antigenic protein specific to non-A non-B hepatitis.

EXAMPLE 2: Preparation of cDNA Containing the Full Length Gene Sequence

Messenger RNA was prepared as described in Example 1 and cDNA was synthesized using Okayama vector according to the conventional method described in Molecular Cloning, p. 211 et seq. The procedures used to synthesize cDNA were as follows:

To 300 µl of a solution (10 mM Tris-HCl, pH 7.5, 6 mM MgCl₂, 10 mM NaCl), there were added 400 µg of

pCDV 1 (Okayama and Berg, Mol. Cell. Biol., 3, 280 (1983)) and 500 units of KpnI (TAKARA SHUZO, Japan), all restriction enzymes used hereinafter having been manufactured by TAKARA SHUZO (Japan) unless otherwise noted. Reaction was effected at 37°C for 6 hours to cut the plasmid at KpnI site therein. After phenol-chloroform extraction, ethanol precipitation was effected to recover DNA.

The DNA (about 200 µg) cleaved with Kpnl was added to 200 µl of a solution which was obtained by adding dTTP in a concentration of 0.25 mM to a buffer (TdT buffer): 40 mM sodium cacodylate, 30 mM Tris-HCl (pH 6.8), 1 mM CaCl₂, 0.1 mM dithiothreitol (DTT). Further, 81 units of terminal deoxynucleotidyl transferase (TdT, manufactured by P-L Biochemicals) was also added. Reaction was effected at 37°C for 11 minutes. Thus, a poly(dT) chain (about 67 deoxythymidylic acid residues) was added to the 3' end at the Kpnl-cleaved site of pCDV 1. After phenol-chloroform extraction and ethanol-precipitation, about 100 µg of pCDV 1 DNA to which poly(dT) chain had been added was recovered from the reaction mixture.

The thus obtained DNA was added to 150 µl of a buffer (10 mM Tris-HCl, pH 7.5, 6 mM MgCl₂, 100 mM NaCl), and <u>Hpal</u> (360 units) was also added, followed by reaction at 37°C for 2 hours. The reaction mixture was subjected to electrophoresis on agarose gel to separate and recover about 3.1 Kbp DNA fragment. Thus, there was obtained about 60 µg of poly(dT)-containing pCDV 1.

The thus obtained DNA was dissolved in 500 µl of a solution (10 mM Tris-HCl, pH 8.0, 1 mM EDTA), incubated at 65°C for 5 minutes, and cooled on ice. After adding 5 M NaCl (50 µl), the mixture was subjected to chromatography on oligo(dA) cellulose column (Colaborative Research). DNA having a poly(dT) chain of sufficient length was adsorbed on the column and eluted with a solution: 10 mM Tris-HCl (pH 8.0), 1 mM EDTA. Thus, there was obtained 27 µg of pCDV 1 to which poly(dT) chain had been added, abbreviated hereinafter as vector primer.

A linker DNA was prepared in the following manner: To 200 µl of a solution (10 mM Tris-HCl, pH 7.5, 6 mM MgCl₂, 50 mM NaCl), there were added about 14 µg of pL 1 (Okayama and Berg, Mol. Cell. Biol., 3, 280 (1983)) and 50 units of Pstl. Reaction was effected at 37°C for 4 hours to cut the pL 1 DNA at Pstl site. Phenol-chloroform extraction and ethanol precipitation of the reaction product gave about 13 µg of pL 1 DNA cleaved at Pstl site.

25

*3*0

35

40

65

The thus obtained DNA (about 13 µg) was added to 50 µl of the TdT buffer containing dGTP at a final concentration of 0.25 mM, and 54 units of TdT (P-L Biochemicals) was also added. The mixture was incubated at 37°C for 13 minutes to add a (dG) chain (about 14 deoxyguanylic acid residues) to the 3′ end at the Pstl-cleaved site of pL 1. After phenol-chloroform extraction, DNA was recovered by ethanol precipitation. The thus obtained DNA was added to 100 µl of a buffer (10 mM Tris-HCl, pH 7.5, 6 mM MgCl₂, 60 mM NaCl), and 80 units of Hindlil was also added. The mixture was incubated at 37°C for 3 hours to cut the pL 1 DNA at Hindlil site. The reaction product was fractionated by agarose gel electrophoresis. About 0.5 Kb DNA fragment was recovered by the DEAE paper method: Dretzen et al., Anal. Biochem., 112, 295 (1981). Thus, there was

obtained an oligo(dG) chain-containing linker DNA, hereinafter abbreviated simply as linker DNA. The aforementioned poly(A)+ RNA (about 2 μg) prepared in the same manner as in Example 1 and the vector primer (about 1.4 μg) were dissolved in 22.3 μl of a solution: 50 mM Tris-HCl (pH 8.3), 8 mM MgCl₂, 30 mM KCl, 0.3 mM DTT, 2 mM dNTP (dATP, dTTP, dGTP and dCTP) and 10 units of ribonuclease inhibitor (P-L Biochemicals). To the solution, there was added 10 units of reverse transcriptase manufactured by SEIKAGAKU KOGYO (Japan). Incubation was effected at 37°C for 40 minutes to synthesize a DNA complementary to the mRNA. After phenol-chloroform extraction and ethanol precipitation, the vector primer DNA to which a double stranded RNA-DNA had been added was recovered.

The thus obtained vector primer DNA containing RNA-DNA double stranded chain was dissolved in 20 µl of TdT buffer containing 60 µM dCTP and 0.2 µg poly(A). After adding 14 units of TdT (P-L Biochemical), the mixture was incubated at 37°C for 8 hours to add a (dC) chain of 12 deoxycytidylic acid residues to the 3' end of the cDNA. The reaction product was extracted with phenol-chloroform and precipitated with ethanol to recover a cDNA-vector primer DNA to which a (dC) chain had been added.

The thus obtained (dC) chain-containing cDNA-vector primer DNA was dissolved in 400 μ i of a solution (10 mM Tris-HCl, pH 7.5, 6 mM MgCl₂, 60 mM NaCl), and 20 units of Hindlii was also added. The mixture was incubated at 37°C for 2 hours to cut the DNA at Hindlii site. The reaction product was extracted with phenol-chloroform and precipitated with ethanol. Thus, there was obtained 0.5 pmole of a (dC) chain-containing cDNA-vector primer DNA.

The thus obtained (dC) chain-containing cDNA-vector primer DNA (0.08 pmole) and the aforementioned linker DNA (0.16 pmole) were dissolved in 40 μl of a solution: 10 mM Tris-HCl (pH 7.5), 0.1 M NaCl, 1 mM EDTA. The resulting solution was incubated at 65°C for 10 minutes, at 42°C for 25 minutes, and then at 0°C for 30 minutes. The reaction mixture was adjusted to 20 mM Tris-HCl (pH 7.5), 4 mM MgCl₂, 10 mM (NH₄)₂SO₄, 0.1 M KCl and 0.1 mM β-NAD in a total volume of 400 μl.

To the reaction mixture, there was added 10 units of <u>E. coli</u> DNA ligase (New England Biolabs), followed by incubation overnight at 11°C. After adjusting the concentrations of dNTP and β-NAD in the reaction mixture to 40 μM and 0.15 mM, respectively, by supplementing necessary reagents, 5 units of <u>E. coli</u> DNA polymerase I (P-L Biochemicals) and 2 units of <u>E. coli</u> ribonuclease H (P-L Biochemicals) were added to the reaction mixture. The mixture was incubated at 12°C for one hour and then at 25°C for one hour.

In the course of the above reactions, a recombinant DNA containing the cDNA was cyclized and the RNA portion of the RNA-DNA double stranded chain was substituted by DNA. Thus, a desired recombinant plasmid containing a complete double-stranded DNA was produced.

The recombinant plasmid was used to transform competent cells of E. coli strain MC1064 prepared by conventional methods. Approximately 50,000 transformants were fixed on a nitrocellulose filter. These colonies were screened according to the colony hybridization method described in Molecular Cloning, Cold Spring Harbor Laboratory, p. 329 et seq. (1982) using the cDNA fragment obtained in Example 1 as a 32P-labelled probe. Thus, three clones showed strong hybridization at 42°C.

These positive clones were analyzed in detail by Southern method: J. Mol. Biol., 98, 503 (1975). There was obtained the desired full length cDNA of a gene coding for an antigenic protein specific to non-A non-B hepatitis. The base sequence of the cDNA is shown in Fig. 4.

5

10

15

20

25

30

35

55

65

The expression vector containing the full length cDNA was designated as pCDVCL-I.

EXAMPLE 3: Preparation of Expression Vector and Transformant and Expression of Specific Antigen

A. Preparation of Expression Vector and Transformant

I) Modification of N-terminus (Fig. 5):

i) in 100 μl of a buffer (10 mM Tris-HCl, pH 7.5, 100 mM NaCl, 6 mM MgCl₂), pCDVCL-I (5 μg) was digested with Pvul (10 units) at 37°C for 2 hours. The reaction mixture was heated at 75°C for 15 minutes to deactivate the enzyme, dialyzed against water, and dried. The cleaved plasmid DNA was treated with T4 DNA polymerase (4 units) in 40 μl of a system: 33 mM Tris-acetic acid (pH 7.9), 66 mM potassium acetate, 10 mM magnesium acetate and 0.5 mM dithiothreitol, to which 2 mM 4-deoxytrlphosphate had been added; thus, the 3' protruding end of the plasmid DNA was filled in to produce a blunt end. The thus treated mixture was heated at 70°C for 10 minutes to deactivate the enzyme, dialyzed against water, and dried. The thus obtained plasmid DNA was then stored in the form of an aqueous solution (50 μl). This plasmid DNA fragment is hereinafter designated as Fragment I.

ii) On the other hand, pCDVCL-I (20 μg) was digested with Ncol and Hindill (each 20 units) at 37°C for 2 hours in 100 μl of a buffer: 10 mM Tris-HCl (pH 7.5), 100 mM NaCl, 6 mM MgCl₂. The plasmid DNA was subjected to 5% acrylamide gel electrophoresis at 10 V/cm for 1.5 hours in a buffer: 89 mM Tris, 89 mM boric acid, 2 mM EDTA. The gel was stained with 0.05% aqueous ethidium bromide solution and two gel slices corresponding to DNA fragments of larger molecular weights were excised from the gel under ultraviolet radiation at 340 nm. The gel slices were crushed by means of a glass rod, suspended into 4 ml of a buffer for DNA extraction (0.5 M ammonium acetate, 10 mM magnesium acetate, 1 mM EDTA, 0.1% sodium laurylsulfate), and allowed to stand overnight at 37°C to extract DNA from the gel. The materials were subjected to centrifugation at 10,000 rpm for 15 minutes to eliminate larger gel pieces, and passed through a glass filter to remove smaller gel pieces. The DNA was purified by effecting ethanol percipitation three times and stored in the form of an aqueous solution (200 μl). This plasmid DNA fragment is hereinafter designated as Fragment II.

iii) A primer of the DNA portion to be modified as shown below (51 bases) was synthesized by a DNA synthesizer, NiKKAKI (Japan), Applied Biosystem MODEL 380A. The synthesized DNA was overnight reacted with concentrated aqueous ammonia at 55°C to deprotect and purified by reversed HPLC before use.

The synthetic primer (150 pmole) was treated with T4 polynucleotide kinase (20 units) in 10 μ l of a kinase buffer (50 mM Tris-HCl, pH 8.0, 10 mM MgCl₂, 5 mM dithiothreitol) to phosphorylate the 5' end thereof.

iv) Fragment I (0.05 pmole), Fragment II (0.05 pmole) and 5'-phosphorylated primer (45 pmole) were added to 12 μl of 5 x polymerase-ligase buffer (0.5 M NaCl, 32.5 mM Tris-HCl, pH 7.5, 40 mM MgCl₂, 5 mM β-mercaptoethanol) to make the total volume of the mixture 34.8 μl. The mixture was boiled at 100°C for 3 minutes, immediately after which it was placed in a thermostat at 30°C and allowed to stand for 30 minutes. The mixture was allowed to stand at 4°C for 30 minutes and then on ice for 10 minutes to form a heteroduplex.

To an aqueous solution (11.6 µl) containing the heteroduplex, there were added 2.5 mM 4-deoxynucleotide

9

triphosphate (2 μ), 10 mM ATP (2 μ), Klenow enzyme (2 units) and T4 DNA ligase (0.5 units) to form a mixture of 20 μ l in total volume. The mixture was reacted overnight at 16°C to cyclize the DNA.

An aqueous solution (2 µl) containing the circular DNA was used to transform E. coll HB101 strain according to conventional methods. Plasmids were separated from the transformant and purified in conventional manners. The plasmid was cleaved with restriction enzyme Hindill and subjected to 5% acrylamide gel electrophoresis. Thus, two separate fragments were collected as desired modified, variant plasmids. Since resulting variant plasmids might often be admixed with original wild-type plasmids, the thus obtained variant plasmids were again employed to transform E. coll HB101 so as to purify the plasmid.

Thus, a purified plasmid pCV44H was obtained (Fig. 5).

II) Modification of C-terminus (Fig. 6):

- i) Plasmid pCDVCL-I (5 μg) was treated in the same manner as in I) i) described above to produce Fragment I.
- ii) Plasmid pCDVCL-I (20 μg) was treated in the same manner as in I) ii) described above except that Ncol and Nsii (each 5 units) were employed. Thus, Fragment II was produced.
- III) In the same manner as in I) III) described above, the following primer (46 bases) was synthesized and the 5' end thereof was phosphorylated.

20

35

40

45

50

55

80

10

15

Primer GCACAAGGAAAAAATGAGATCTGTCGACGGTTCACGTA

xxxxxxxxxxxxx

25 (Original sequence) (------AGATATGTGAA*A----
AATTTCC wherein x represents a base substitution and * represents ------) an addition.

iv) The Fragment I and II and the 5'-phosphorylated primer obtained above in II) i) to iii) were treated in the same manner as in I) iv) described above. Thus, plasmid pCV44B was obtained (Fig. 6).

III) Introduction of cDNA coding for specific antigen into expression vector (Fig. 7):

i) In 100 µl of a buffer H (10 mM Tris-HCl, pH 7.5, 100 mM NaCl, 6 mM MgCl₂), 10 µg (about 3 pmole) of pCV44H was cut with Hindill (20 units) and Sacl (20 units) at 37°C for 2 hours. The reaction mixture was subjected to 5% acrylamide gel electrophoresis. Thus, a 467 bp DNA fragment coding for the N-terminus of the specific antigen was separated and purified. This fragment is hereinafter designated as Fragment N.

ii) In 100 μl of the buffer H, 10 μg (about 3 pmole) of pCV448 was cleaved with Bglll (20 units) and Saci (20 units) at 37°C for 2 hours. The reaction mixture was subjected to 5% acrylamide gel electrophoresis to isolate and purify a 836 bp DNA fragment coding for the C-terminus of the specific antigen. The thus obtained fragment is hereinafter designated as Fragment C.

iii) In 20 μl of buffer H, 2 μg (about 1 pmole) of an expression vector pUSΔH was cut with Hindlii (2 units) and Bglii (2 units) at 37°C for 2 hours. The reaction mixture was extracted with an equal volume of water-saturated phenol to remove proteins. After extracting the phenol with ether, the reaction mixture was dialyzed against water to desalt, and concentrated by a vacuum pump. Thus, there was obtained 10 μl of an aqueous solution containing an expression vector fragment HB.

iv) Fragment N (0.5 pmole), Fragment C (0.5 pmole) and the expression vector fragment HB (0.1 pmole) were mixed and reacted with T4 DNA ilgase (1 unit) at 4°C for 16 hours in 10 μl of a buffer (10 mM Tris-HCl, pH 7.5, 1 mM dithiothreitol, 6 mM MgCl₂, 1 mM ATP). The reaction mixture (3 μl) was used to transform commercially available E. coli JM109 competent cell according to conventional methods. The resulting transformants were selected in L broth plate (bactopeptone 10 g, yeast exstract 5 g, NaCl 10 g, agar 15 g per liter) containing 20 μg/ml ampicillin. Thus, there was obtained an expression vector pCZ44 containing the specific antigen gene inserted thereinto (Fig. 7).

B. Expression of Specific Antigen

E. coli strain JM109 possessing pCZ44 was cultured overnight at 30°C in L broth. The culture was inoculated in a fresh L broth with a dilution of 1/50 and cultured with shaking at 30°C for 2 hours. After IPTG (isopropylthio-β-D-galactopyranoside) was added to the medium in a concentration of 2 mM, shaking culture was continued at 30°C for further 3 hours. The cells were collected by centrifugation at 6,500 rpm for 10 minutes and suspended in a buffer (0.9% NaCl, 10 mM Tris-HCl, pH 7.5) to store.

65

£

C. Verification of Expression of Specific Antigen

The thus obtained cell culture (0.3 ml) was subjected to 10% SDS polyacrylamide gel electrophoresis at 120 V for one hour in a buffer (Tris 3g/1, glycine 14.4 g/1, 0.1% SDS). The gel was removed, placed on a nitrocellulose filter, interposed between filter papers and electrophoresed at 5 V/cm, 4°C in a buffer (Tris 3g/1, glycine 14.4 g/1) to transferred proteins in the gel onto the nitrocellulose filter. The nitrocellulose filter was rinsed with TBS buffer (10 mM Tris-HCl, pH 7.5, 50 mM NaCl), immersed into 400 ml of TBS buffer containing 3% gelatine and shaked at 40°C for one hour to block the nitrocellulose filter.

To TBS buffer containing 1% gelatine, there was added a monoclonal antibody directed to a non-A non-B hepatitis-specific antigen (OD₂₈₀ = 4.3) with a dilution of 1/400. The resulting mixture and the nitrocellulose filter were put into a vinyl bag so that the mixture was present in an amount of 2 ml per filter. Reaction was effected at room temperature for 16 hours. The reaction mixture was washed three times with 400 ml of TBS buffer containing 0.05% Tween 20 for 10 minutes.

To TBS buffer containing 1% gelatine, there was added a labelled secondary antibody, anti-mouse IgG-PAP (horseradish peroxidase, Bio Rad), with a dilution of 1/1000. The resulting mixture and the nitrocellulose filter were put into a vinyl bag so that the mixture was present in an amount of 2 ml per filter. Reaction was effected at room temperature for 2 hours. The reaction mixture was washed three times with 400 ml of TBS buffer containing 0.05% Tween 20 for 10 minutes.

Color formation was effected by immersing the filter into 20 ml of TBS buffer containing 12 mg of 4-chloro-1-naphthol (Bio Rad) and hydrogen peroxide. After completion of color formation, the filter was thoroughly washed with water, put into a vinyl bag containing water, and stored in a dark and cold place. Such a test effected showed that a protein reacting with

The monoclonal antibody was found at the same position (44 Kd) as found in the case of the specific antigen derived from infected chimpanzee liver. This verifies that such a specific antigen can be in fact expressed in Ecoli. The invention thus also relates to a process for the in vitro diagnosis of NON-A NON-B hepatitis, which comprises contacting a liver sample and/or a serum sample taken from a patient possibly infected with a NON-A NON-B hepatitis with the protein whose formula appears in claim 3 hereafter or a part thereof for a time and under conditions sufficient to allow for the production of a complex between said protein or part thereof with the antibodies contained in the patient sample and detecting the presence of the immunologic complex, particularly when the patient is suffering from NON-A NON-B hepatitis.

Any part of said protein, or any recombinant, produced by genetic engineering and including the aminoacid sequence of said protein or part of said protein can be substituted for above-said protein, it being understood that the said recombinant protein or part of said protein are specifically recognized by the same antibodies as those which recognize said protein.

In other words the invention relates to all recombinant proteins or protein fragments which bind to antibodies contained in a liver extract or serum sample, or both, and originating from a patient suffering from NON-A NON-B hepatitis.

The invention also relates to a process for detecting in vitro an infection by a NON-A NON-B hepatitis virus, which process comprises contacting the DNA of claim 4, or a fragment thereof, under suitable hybridization conditions, with a sample of liver extract and/or serum sample originating from the patient to be diagnosed and in which the nucleic acid components had previously been made accessible to hybridization, to form a hybridization product between said DNA of claim (probe) and the viral DNA of a NON-A NON-B hepatitis B virus, and detecting said hybridization product, particularly in the case where the patient is Indeed infected with a NON-A NON-B virus.

Claims

- 1. A DNA fragment which contains a base sequence coding for a non-A non-B hepatitis-specific antigenic protein occurring in cells of the liver affected with non-A non-B hepatitis.
- 2. The DNA fragment in accordance with claim 1, in which the cells of the liver are derived from a human or chimpanzee individual.
- 3. The DNA fragment in accordance with claim 1, in which the antigenic protein specific to non-A non-B hepatitis has the whole or a part of the amino acid sequence represented by the formula:

65

5

10

15

20

25

30

35

45

50

55

60

10 Met Ala Val Thr Thr Arg Leu Thr Trp Leu His Glu Lys Ile Leu 30 20 Gln Asn His Phe Gly Gly Lys Arg Leu Ser Leu Leu Tyr Lys Gly 40 Ser Val His Gly Phe His Asn Gly Val Leu Leu Asp Arg Cys Cys 60 50 Asn Gln Gly Pro Thr Leu Thr Val Ile Tyr Ser Glu Asp His Ile Ile Gly Ala Tyr Ala Glu Glu Gly Tyr Gln Glu Arg Lys Tyr Ala 90 80 Ser Ile Ile Leu Phe Ala Leu Gln Glu Thr Lys Ile Ser Glu Trp 100 Lys Leu Gly Leu Tyr Thr Pro Glu Thr Leu Phe Cys Cys Asp Val 120 110 Ala Lys Tyr Asn Ser Pro Thr Asn Phe Gln Ile Asp Gly Arg Asn

130 Arg Lys Val Ile Met Asp Leu Lys Thr Met Glu Asn Leu Gly Leu 140 150 Ala Gln Asn Cys Thr Ile Ser Ile Gln Asp Tyr Glu Val Phe Arg 160 Cys Glu Asp Ser Leu Asp Glu Arg Lys Ile Lys Gly Val Ile Glu 170 180 Leu Arg Lys Ser Leu Leu Ser Ala Leu Arg Thr Tyr Glu Pro Tyr 190 Gly Ser Leu Val Gln Gln Ile Arg Ile Leu Leu Gly Pro Ile 200 210 Gly Ala Gly Lys Ser Ser Phe Phe Asn Ser Val Arg Ser Val Phe 220 Gln Gly His Val Thr His Gln Ala Leu Val Gly Thr Asn Thr Thr 230 240 Gly Ile Ser Glu Lys Tyr Arg Thr Tyr Ser Ile Arg Asp Gly Lys 250 Asp Gly Lys Tyr Leu Pro Phe Ile Leu Cys Asp Ser Leu Gly Leu 260 270 Ser Glu Lys Glu Gly Gly Leu Cys Met Asp Asp Ile Ser Tyr Ile 280 Leu Asn Gly Asn Ile Arg Asp Arg Tyr Gln Phe Asn Pro Met Glu 290 300 Ser Ile Lys Leu Asn His His Asp Tyr Ile Asp Ser Pro Ser Leu 310 Lys Asp Arg Ile His Cys Val Ala Phe Val Phe Asp Ala Ser Ser 320 330 Ile Glu Tyr Phe Ser Ser Gln Met Ile Val Lys Ile Lys Arg Ile 340 Arg Arg Glu Leu Val Asn Ala Gly Val Val His Val Ala Leu Leu 350 360 Thr His Val Asp Ser Met Asp Leu Ile Thr Lys Gly Asp Leu Ile 370 Glu Ile Glu Arg Cys Val Pro Val Arg Ser Lys Leu Glu Glu Val 380 390 Gln Arg Lys Leu Gly Phe Ala Leu Ser Asp Ile Ser Val Val Ser

Asn Tyr Ser Ser Glu Trp Glu Leu Asp Pro Val Lys Asp Val Leu Leu Ser Ala 410 Leu Arg Met Leu Trp Ala Ala Asp Asp Phe Leu Glu Asp Leu Pro Phe Glu Gln Ile Gly Asn Leu Arg Glu Glu Glu Ile Ile Asn Cys Ala Gln Gly Lys Lys ***.

4. The DNA fragment in accordance with claim 1, in which the base sequence comprises the whole or a part of the base sequence represented by the formula:

			1	10			20			30			4	0
5'	ATG	GCA												
3 1														
									-	20			9.0	
		~~							CCC					መልመ
	CTG	CAA	AAT	CAT	TTT	GGA		AAG	CGG	C11	AGC	CII		INI
		90		-	10	00		1	110			120		
									AAT				CTT	GAC
•														
								1				•		
									-				m a m	хсш
	AGA	TGT	TGT	AAT	CAA	GGG	CCT	ACT	CTA	ACA	GTG	ATT	TAT	AGT
1	L70			180			19	90			200			210
	GAA	GAT	CAT	ATT	ATT	GGA	GCA	TAT	GCA	GAA	GAG	GGT	TAC	CAG
			, 											
													•	- 0
	GAA	AGA	AAG	TAT	GCT	TCC	ATC	ATC	CTT	TTT	GCA	CTT	CAA	GAG
										·				
			260			270			21	80			290	
	x cm								GGA				_	GAA
	ACI	AAA	W11	ICA	900		~~~							
		- - -												
		300			3	10		•	320			330		
	ACA	CTG	TTT	TGT	TGT	GAC	GTT	GCA	AAA	TAT	AAC	TCC	CCA	ACT

34	10		3	350			360			37	7.0		
	TTC											ATG	GAC
380			390			40	0		4	110			420
TTA	AAG	ACA	ATG	GAA	AAT	CTT	GGA	CTT	GCT	CAA	AAT	TGT	ACT
ATC	TCT	ATT	CAG	GAT	TAT	GAA	GTT	TTT	CGA	TGC	GAA	GAT	TCA
								'-					
											_		
								49					
CTG	GAC	GAA	AGA	AAG	ATA	AAA	GGG	GTC	ATT	GAG	CTC	AGG	AAG
	510			5 -	20			530			E 4 O		
3.00	TTA											CCX	mcc
AGC	TIM		101		110	AGA	ACI	IMI	GAA	CCA	IAT	GGA	100
55	50 -			560			570			58	30		
	GTT											ATT	GGA
590			600			61	LO		(620			630
GCT	GGG	AAG	TCT	AGC	TTT	TTC	AAC	TCA	GTG	AGG	TCT	GTT	TTC
		_											
CAA	GGG	CAT	GTA	ACG	CAT	CAG	GCT	TTG	GTG	GGC	ACT	AAT	ACA
					. — — —								
	•	SBA			690			70	۱0			710	
ACT	GGG												GAC
	720			7:	30		•	740			750		
GGG	AAA	GAT	GGC	AAA	TAC	CTG	CCA	TTT	ATT	CTG	TGT	GAC	TCA
			_							_		•	
	60												
CTG	GGG	CTG	AGT	GAG	AAA	GAA	GGC	GGC	CTG	TGC	ATG	GAT	GAC
800			910			٥٠	20			220			940
ATA													
WIW	100	IWC	wir.	110	MMC	201	77.	WII	CQ1	GW1	AUA	TWC	CAG
		~ 											
		я	50		1	860			870			81	30
TTT	AAT												

		8	390			900			91	LO		9	20	
	ATT	GAT	TCC	CCA	TCG	CTG	AAG	GAC	AGA	ATT	CAT	TGT	GTG	GCA
		930			94	10		9	950			960	•	
	TTT	GTA	TTT	GAT	GCC	AGC	TCT	ATT	GAA	TAC	TTC	TCC	TCT	CAG
	97	70		9	980			990			100	00		
	ATG	ATA	GTA	AAG	ATC	AAA	AGA	ATT	CGA	AGG	GAG	TTG	GTA	AAC
10	10			1020			103	30		10	140		1	L050
	GCT	GGT	GTG	GTA	CAT	GTG	GCT	TTG	CTC	ACT	CAT	GTG	GAT	AGC
			10	50		10	70		1	1080			109	90
	ATG	GAT	CTG	ATT	ACA	AAA	GGT	GAC	CTT	ATA	GAA	ATA	GAG	AGA
		13	100		1	1110			112	20		11	130	
	TGT	GTG	CCT	GTG	AGG	TCC	AAG	CTA	GAG	GAA	GTC	CAA	AGA	AAA
								~~~						
	1	1140			115	50		13	160		]	L170		
	CTT	GGA	TTT	GCT	CTT	TCT	GAC	ATC	TCG	GTG	GTT	AGC	AAT	TAT
										~~~				
	13	180		1:	190		•	1200			12	LO		
	TCC	TCT	GAG	TGG	GAG	CTG	GAC	CCT	GTA	AAG	GAT	GTT	CTA	ATT
12	220		•	1230			124	40		13	250			1260
	CTT	TCT	GCT	CTG	AGA	CGA	ATG	CTA	TGG	GCT	GCA	GAT	GAC	TTC
						~								
	TTA	GAG	GAT	TTG	CCT	TTT	GAG	CAA	ATA	GGG	AAT	CTA	AGG	GAG
		13	310			1320			13	30				
	GAA	ATT	ATC	AAC	TGT	GCA	CAA	GGA	AAA	AAA				
											5 *			

wherein the sign "-" represents a base complementary to the base shwon just above each sign.

- 5. An expression vector in which a DNA fragment containing a base sequence coding for a non-A non-B hepatitis-specific antigen is intoduced into a cloning site present downstream from a promoter of said vector.
- 6. The expression vector in accordance with claim 5, in which the promoter is controllable by a regulatory factor.

5

10

15

20

25

30

35

*5*0

55

60

- 7. The expression vector in accordance with claim 5, in which the promoter operates in a microorganism.
 - 8. The expression vector in accordance with claim 5, in which the promoter operates in an eukaryote.
- 9. A transformant obtained by transforming a host with an expression vector in which a DNA fragment containing a base sequence coding for a non-A non-B hepatitis-specific antigen is introduced into a cloning site present downstream from a promoter of said vector.
- 10. The transformant in accordance with claim 9, in which the host is Escherichia coli or Bacillus subtilis.
- 11. A process for producing an antigen specific to non-A non-B hepatitis comprising introducing a DNA fragment containing a base sequence coding for said specific antigen into a cloning site present downstream from a promoter of a vector for expression, introducing the expression vector containing said DNA fragment into a host, culturing said transformed host, and collecting the produced and accumulated antigen.

17

Fig. la

5' ATG GCA GTG ACA ACT CGT TTG ACA TGG TTG CAT GAA AAG ATC CTG CAA AAT CAT TTT GGA GGG AAG CGG CTT AGC CTT CTC TAT AAG GGT AGT GTC CAT GGA TTC CAT AAT GGA GTT TTG 130 140 CTT GAC AGA TGT TGT AAT CAA GGG CCT ACT CTA ACA GTG ATT TAT AGT GAA GAT CAT ATT ATT GGA GCA TAT GCA GAA GAG GGT TAC CAG GAA AGA AAG TAT GCT TCC ATC ATC CTT TTT GCA CTT CAA GAG ACT AAA ATT TCA GAA TGG 280 290 AAA CTA GGA CTA TAT ACA CCA GAA ACA CTG

Fig. 1b

TTT TGT TGT GAC GTT GCA AAA TAT AAC TCC CCA ACT AAT TTC CAG ATA GAT GGA AGA AAT AGA AAA GTG ATT ATG GAC TTA AAG ACA ATG GAA AAT CTT GGA CTT GCT CAA AAT TGT ACT ATC TCT ATT CAG GAT TAT GAA GTT TTT CGA 470 480 TGC GAA GAT TCA CTG GAC GAA AGA. AAG ATA AAA GGG GTC ATT GAG CTC AGG AAG AGC TTA CTG TCT GCC TTG AGA ACT TAT GAA CCA TAT GGA TCC CTG GTT CAA CAA ATA CGA ATT CTG CTG CTG GGT CCA ATT GGA GCT GGG AAG TCT

Fig. 1c

AGC TTT TTC AAC TCA GTG AGG TCT GTT TTC 650 · CAA GGG CAT GTA ACG CAT CAG GCT TTG GTG 680 · GGC ACT AAT ACA ACT GGG ATA TCT GAG AAG TAT AGG ACA TAC TCT ATT AGA GAC GGG AAA GAT GGC AAA TAC CTG CCA TTT ATT CTG TGT GAC TCA CTG GGG CTG AGT GAG AAA GAA GGC GGC CTG TGC ATG GAT GAC ATA TCC TAC ATC TTG AAC GGT AAC ATT CGT GAT AGA TAC CAG 850 860 TTT AAT CCC ATG GAA TCA ATC AAA TTA AAT 880 890 CAT CAT GAC TAC ATT GAT TCC CCA TCG CTG

Fig. 1d

910 920 930 AAG GAC AGA ATT CAT TGT GTG GCA TTT GTA

940 950 960 TTT GAT GCC AGC TCT ATT GAA TAC TTC TCC

970 980 . 990 TCT CAG ATG ATA GTA AAG ATC AAA AGA ATT

1000 1010 1020 CGA AGG GAG TTG GTA AAC GCT GGT GTG GTA

1030 1040 1050 CAT GTG GCT TTG CTC ACT CAT GTG GAT AGC

1060 1070 1080
ATG GAT CTG ATT ACA AAA GGT GAC CTT ATA

1090 1100 1110
GAA ATA GAG AGA TGT GTG CCT GTG AGG TCC

1120 1130 1140

AAG CTA GAG GAA GTC CAA AGA AAA CTT GGA

1150 1160 1170 TTT GCT CTT TCT GAC ATC TCG GTG GTT AGC

1180 1190 1200
AAT TAT TCC TCT GAG TGG GAG CTG GAC CCT

Fig. 2

AAAAATTTATTTGCTTTCAGGAAAATTTTTCTGT TTTTTAAATAAACGAAAGTCCTTTTAAAAAGACA

ATAATGTGTGGAATTGTGAGCGGATAACAATTTC TATTACACACCTTAACACTCGCCTATTGTTAAAG

...;

.

Fig. le

1210 1220 1230 GTA AAG GAT GTT CTA ATT CTT TCT GCT CTG

1240 1250 1260

AGA CGA ATG CTA TGG GCT GCA GAT GAC TTC

1270 1280 1290 TTA GAG GAT TTG CCT TTT GAG CAA ATA GGG

1300 1310 1320
AAT CTA AGG GAG GAA ATT ATC AAC TGT GCA

1330 CAA GGA AAA AAA 3' 5'

300 CCA Pro	360 GAA Glu	420 TGC Cys	480 CTG Leu	540 CTG Leu	600 CAA Gln
TCC	ATG Met	CGA	TTA Leu	CTG	TTC
AAC	ACA	TTT Phe	AGC	ATT Ile	GTT Val
		410 GTT Val	470 AAG Lys	530 ATA CGA Ile Arg	590 AGG TCT Arg Ser
290 AAA TAT Lys Tyr	350 TTA AAG Leu Lys	410 GAA GTT Glu Val	470 AGG AAG Arg Lys	ATA Ile	590 AGG TCT Arg Ser
GCA	GAC	tat Tyf	CTC	CAA Gln	GTG Val
io GTT Val	340 ATT ATG Ile Met	400 CAG GAT Gln Asp	460 ATT GAG Ile Glu	520 GTT CAA Val Gln	580 AAC TCA Asn Ser
280 GAC GTT Asp Val	340 ATT ATG Ile Met	400 CAG GAT Gln Asp	460 ATT G Ile G	520 GTT C Val G	580 AAC T Asn S
TGT Cys	GTG Val	ATT Ile	GTC Val	CTG	TTC
TGT Cys	AAA Lyb	TCT	666 61y	TCC	TTT Phe
270 TTT Phe	330 AGA Arg	390 ATC Ile	450 AAA Lys	510 GGA Gly	570 AGC Ser
CTG Leu		ACT	AAG ATA Lys Ile	TAT	
ACA	AGA AAT Arg Asn	TGT Cys	AAG	CCA	AAG Lys
260 CCA GAA Pro Glu	320 GAT GGA ASP Gly	380 CAA AAT Gln Asn	440 GAA AGA Glu Arg	500 TAT GAA TYF Glu	550 GGT CCA ATT GGA GCT GGG AAG TCT Gly Pro Ile Gly Ala Gly Lys Ser
CCA Ord		CAA Gln	GAA Glu	TAT Tyr	GCT Ala
ACA	ATA Ile	GCT	430 TCA CTG GAC GAA AGA Ser Leu Asp Glu Arg	ACT	550 CCA ATT GGA Pro Ile Gly
250 CTA TAT Leu Tyr	310 TTC CAG Phe Gln	370 GGA CTT Gly Leu	CTG Leu Leu	490 TTG AGA Leu Arg	550 A ATT
	31 TTC Phe				SCA Pro
gga gly	AAT	CTT	gat Asp	GCC Ala	
CTA	ACT	AAT	GAA	TCT	CIG

660 TAT TYT	720 GAC Asp	780 TTG Leu	840 CAT His	900 TTT Phe	960 CGA Arg	1020 ATG Met
AAG Lys	TGT Cys	ATC Ile	AAT	GTA Val	ATT Ile	AGC Ser
GAG Glu	CTG	TAC	TTA	TTT Phe	AGA Arg	gat Asp
650 TCT Ser	710 ATT Ile	770 TCC Ser	830 AAA Lys	890 GCA Ala	950 AAA Lys	1010 T GTG s Val
650 ATA TCT Ile Ser	TTT Phe	ATA Ile	830 ATC AAA Ile Lys	grg Val	ATC Ile	1010 CAT GTG His Val
666 Gly	CCA	GAC Asp	TCA	TGT Cys	AAG Lys	ACT
640 ACA ACT Thr Thr	700 TAC CTG TYr Leu	760 ATG GAT Met Asp	820 ATG GAA Met Glu	30 CAT His	10 GTA Val	1000 TTG CTC Leu Leu
640 ACA A Thr	7 TAC TYF	76 ATG Met	87 ATG Met	880 ATT C Ile H	940 ATA G Ile V	1000 TTG C
AAT Asn	AAA Lys	TGC	CCC	AGA	ATG Met	GCT
ACT Thr	GGC G1y	CTG	AAT Asn	GAC	CAG Gln	GTG Val
630 GGC G1y	690 Gat Asp	750 GGC G1y	810 TTT Phe	870 AAG Lys	930 TCT Ser	990 CAT H18
GTG Val	AAA Lys	GGC	CAG Gln	CTG	TCC	GTA Val
TTG	666 61y	GAA Glu	TAC	TCG Ser	TTC Phe	GTG Val
620 CAG GCT Gln Ala	680 AGA GAC Arg Asp	740 GAG AAA Glu Lys	800 GAT AGA Asp Arg	860 CCA Pro	920 GAA TAC Glu Tyr	980 GCT GGT Ala Gly
CAG	AGA	GAG	GAT Asp	TCC Ser	GAA Glu	
CAT	ATT Ile	AGT Ser	CGT	gat Asp	ATT Ile	AAC
610 GTA ACG Val Thr	670 TAC TCT Tyr Ser	730 GGG CTG Gly Leu	790 AAC ATT Asn Ile	850 C ATT r Ile	910 C TCT r Ser	970 TTG GTA Leu Val
		7.3 666 61y	* -	85 TAC	AG Se	_
CAT	ACA Thr	CTG	GGT Gly	GAC	GCC	GAG Glu
GGG G1y	AGG	TCA	AAC Asn	CAT	GAT	AGG

1080 TCC AAG Ser Lys AGG Arg CCT GTG Pro Val 1070 GTG Val AGA TGT Arg Cys 1060 Gla GAG Ile ATA 1050 A GAA Glu ATA Ile Leu CIL GGT GAC 1040 Lys AAA ATT ACA 1030 Asp Leu GAT CTG

1140 AGC AAT Asn Ser GTT Val TCG GTG Ser Val 1130 Ser ATC Ile TCT GAC Ser Asp 1120 CTT Leu GCT Ala GGA TTT Gly Phe 1110 CIT AGA AAA Arg Lys 1100 CAA Glu GAA GTC 1090 GAG Glu Leu CTA

1200 CTG AGA Leu Arg GCT Ala CTT TCT Ser 1190 Leu ATT Ile 1180 GTT CTA Val Leu Asp GAT AAG Lys 1170 CCT GTA Pro Val 1160 GAG CTG GAC Glu Leu Asp Trp \mathbf{TGG} TCC TCT GAG Ser Ser Glu 1150 TAT Tyr

GGG AAT Gly Asn 1260 Gly ATA Ile GAG CAA Glu Gln 1250 TTT Phe TTG CCT Leu Pro 1240 GAT Asp Glu GAG Leu 1230 TTC TTA Phe Asp GAC GCA GAT Ala Asp 1220 GCT CTA TGG Leu Trp 1210 ATG Met CGA Arg

AAA TAG Lvs *** 1300 Lys AAA Lys Gly GGA Gln 1290 GCA CAA Ala TGT Сув ATT ATC AAC Ile Ile Asn 1280 Ile GAG GAA Glu Glu 1270 CTA AGG Leu Arg

80	160	240	320	400	480	560
TCGTTTGACA	TCCATGGATT	CATATTATTG	TAAAATTTCA	CTAATTTCCA	TGTACTATCT	GCTCAGGAAG
AGCAAACTGT	AGGTACCTAA	GTATAATAAC	ATTTTAAAGT	GATTAAAGGT	ACATGATAGA	CGAGTCCTTC
70	150	230	310	390	470	550
CAGTGACAAC	AAGGGTAGTG	TAGTGAAGAT	TTCAAGAGAC	AACTCCCCAA	TGCTCAAAAT	GGGTCATTGA
GTCACTGTTG	TTCCCATCAC	ATCACTTCTA	AAGTTCTCTG	TTGAGGGGTT	ACGAGTTTTA	CCCAGTAACT
60	140	220	300	380	460	540
AGAAGTATGG	CCTTCTCTAT	CAGTGATTTA	CTTTTTGCAC	TGCAAAATAT	ATCTTGGACT	AAGATAAAAG
TCTTCATACC	GGAAGAGATA	GTCACTAAAT	GAAAAACGTG	ACGTTTTATA	TAGAACCTGA	TTCTATTTTC
50	130	210	290	370	450	530
CAACAGATCA	AGCGGCTTAG	CCTACTCTAA	TTCCATCATC	GTTGTGACGT	ACAATGGAAA	GGACGAAAGA
GTTGTCTAGT	TCGCCGAATC	GGATGAGATT	AAGGTAGTAG	CAACACTGCA	TGTTACCTTT	CCTGCTTTCT
40	120	200	280	360	440	520
ACAGACAGTA	TTTGGAGGGA	TAATCAAGGG	GAAAGTATGC	ACACTGTTTT	GGACTTAAAG	AAGATTCACT
TGTCTGTCAT	AAACCTCCCT	ATTAGTTCCC	CTTTCATACG	TGTGACAAAA	CCTGAATTTC	TTCTAAGTGA
30	110	190	270	350	430	510
AGCTCATACT	GCAAAATCAT	ACAGATGTTG	TACCAGGAAA	TACACCAGAA	AAGTGATTAT	TTTCGATGCG
TCGAGTATGA	CGTTTTAGTA	TGTCTACAAC	ATGGTCCTTT	ATGTGGTCTT	TTCACTAATA	AAAGCTACGC
20	100	180	260	340	420	500
CCTCAGCTCT	AAAAGATCCT	GTTTGCTTG	AGAAGAGGGT	TAGGACTATA	AGAAATAGAA	TTATGAAGTT
GGAGTCGAGA	TTTCTAGGA	CAAAACGAAC	TCTTCTCCCA	ATCCTGATAT	TCTTTATCTT	AATACTTCAA
10	90	170	250	330	410	490
GGGGGGCTAC	TGGTTGCATG	CCATAATGGA	GAGCATATGC	GAATGGAAAC	GATAGATGGA	CTATTCAGGA
CCCCCGATG	ACCAACGTAC	GGTATTACCT	CTCGTATACG	CTTACCTTTG	CTATCTACCT	GATAAGTCCT
ສຸຄ						

GATAGCATGG ATCTGATTAC CTATCGTACC TAGACTAATG

AGAATTCGAA GGGAGTTGGT AAACGCTGGT GTGGTACATG TGGCTTTGCT CACTCATGTG
TCTTAAGCTT CCCTCAACCA TTTGCGACCA CACCATGTAC ACCGAAACGA GTGAGTACAC

					•
640	720	800	880	960	1040
TGGGTCCAAT	TTGGTGGGCA	GCCATTTATT	ACGGTAACAT	TCGCTGAAGG	AAAGATCAAA
ACCCAGGTTA	AACCACCCGT	CGGTAAATAA	TGCCATTGTA	AGCGACTTCC	TTTCTAGTTT
630	710	790	870	950	1030
ATTCTGCTGC	GCATCAGGCT	GCAAATACCT	TACATCTTGA	TGATTCCCCA	AGATGATAGT
TAAGACGACG	CGTAGTCCGA	CGTTTATGGA	ATGTAGAACT	ACTAAGGGGT	TCTACTATCA
620	700	780	860	940	1020
ACAAATACGA	GGCATGTAAC	GGGAAAGATG	TGACATATCC	ATGACTACAT	TTCTCCTCTC
TGTTTATGCT	CCGTACATTG	CCCTTTCTAC	ACTGTATAGG	TACTGATGTA	AAGAGGAGAG
610	690	770	850	930	1010
CCCTGGTTCA	GTTTTCCAAG	TATTAGAGAC	TGTGCATGGA	TTAAATCATC	TATTGAATAC
GGGACCAAGT	CAAAAGGTTC	ATAATCTCTG	ACACGTACCT	AATTTAGTAG	ATAACTTATG
600	680	760	840	920	1000
CCATATGGAT	AGTGAGGTCT	GGACATACTC	GAAGGCGGCC	ATCAATCAAA	ATGCCAGCTC
GGTATACCTA	TCACTCCAGA	CCTGTATGAG	CTTCCGCCGG	TAGTTAGTTT	TACGGTCGAG
590	670	750	830	910	990
AACTTATGAA	TTTTCAACTC	GAGAAGTATA	GAGTGAGAAA	ATCCCATGGA	TTTGTATTTG
TTGAATACTT	AAAAGTTGAG	CTCTTCATAT	CTCACTCTTT	TAGGGTACCT	AAACATAAAC
580	660	740	820	900	980
CTGCCTTGAG	AAGTCTAGCT	TGGGATATCT	CACTGGGGCT	TACCAGTTTA	TTGTGTGGCA
GACGGAACTC	TTCAGATCGA	ACCCTATAGA	GTGACCCCGA	ATGGTCAAAT	AACACACCGT
570	650	730	810	890	970
AGCTTACTGT	TGGAGCTGGG	CTAATACAAC	CTGTGTGACT	TCGTGATAGA	ACAGAATTCA
TCGAATGACA	ACCTCGACCC	GATTATGTTG	GACACACTGA	AGCACTATCT	TGTCTTAAGT

1640 ATAATTTCT TATTTAAAGA

CTGTGAAAAA

1610 1620 CTTGGATTTA TGTTCTGTAT GAACCTAAAT ACAAGACATA

1630

1200	1280	1360	1440	1520	1600
CTTGGATTTG	AATTCTTTCT	TAAGGGAGGA	AGATTAAAAT	GATGAAGAAA	AATAATTTTT
GAACCTAAAC	TTAAGAAAGA	ATTCCCTCCT	TCTAATTTTA	CTACTTCTTT	TTATAAAA
1190	1270	1350	1430	1510	1590
CCAAAGAAAA	AGGATGTTCT	ATAGGGAATC	ACATCACAGA	TAATGTCTAG	GAAAAATAAT
GGTTTCTTTT	TCCTACAAGA	TATCCCTTAG	TGTAGTGTCT	ATTACAGATC	CTTTTTATTA
1180	1260	1340	1420	1500	1580
TAGAGGAAGT	GACCCTGTAA	TTTGAGCAA	AAATTTCCTC	TGTGTTTTAT	TCATAATTGT
ATCTCCTTCA	CTGGGACATT	AAAACTCGTT	TTTAAAGGAG	ACACAAAATA	AGTATTAACA
1170	1250	1330	1410	1490	1570
AGGTCCAAGC	GTGGGAGCTG	AGGATTTGCC	AGGTTCACGT	ACCAAAGGGA	CATGATTTAG
TCCAGGTTCG	CACCCTCGAC	TCCTAAACGG	TCCAAGTGCA	TGGTTTCCCT	GTACTAAATC
1160	1240	1320	1400	1480	1560
rgrgccrgrg	ATTCCTCTGA	GACTTCTTAG	GATATGTGAA	AGTAACTAAG	CTAGAAATAA
acacggacac	TAAGGAGACT	CTGAAGAATC	CTATACACTT	TCATTGATTC	GATCTTTATT
1150	1230	1310	1390	1470	1550
TAGAGAGATG	GTTAGCAATT	GGCTGCAGAT	GAAAAAATA	ACCAAAGAGA	TTGTAAATAA
ATCTCTCTAC	CAATCGTTAA	CCGACGTCTA	CTTTTTTAT	TGGTTTCTCT	AACATTTATT
1140	1220	1300	1380	1460	1540
CTTATAGAAA	CATCTCGGTG	GAATGCTATG	TGTGCACAAG	GAAAACACAG	ATTGTAGTAC
GAATATCTTT	GTAGAGCCAC	CTTACGATAC	ACACGTGTTC	CTTTTGTGTC	TAACATCATG
1130	1210	1290	1370	1450	1530
AAAAGGTGAC	CTCTTTCTGA	GCTCTGAGAC	AATTATCAAC	TCAGAAAGGA	TGCATAGAAC
TTTTCCACTG	GAGAAGACT	CGAGACTCTG	TTAATAGTTG	AGTCTTTCCT	ACGTATCTTG

Fig. 5

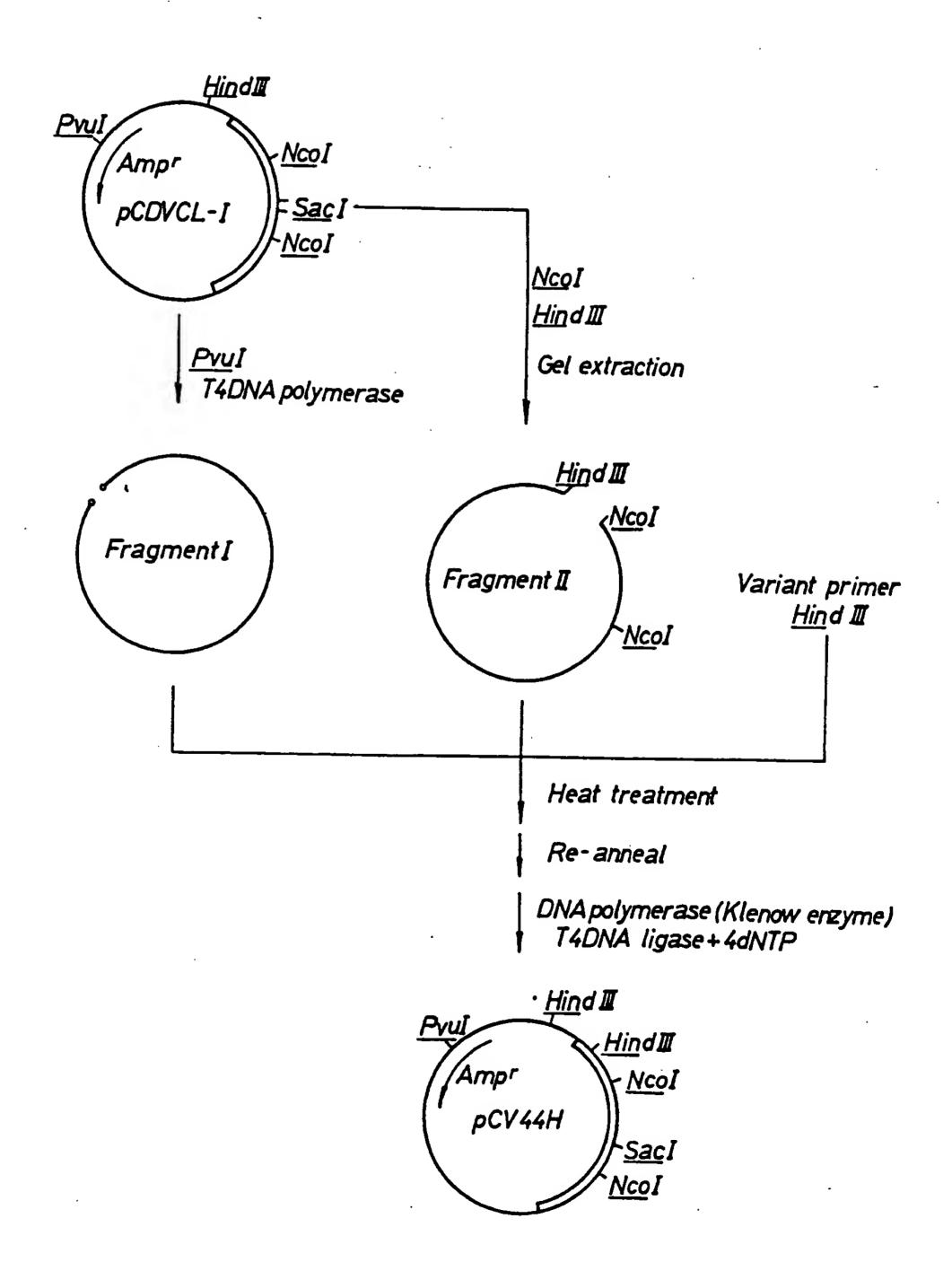


Fig.6

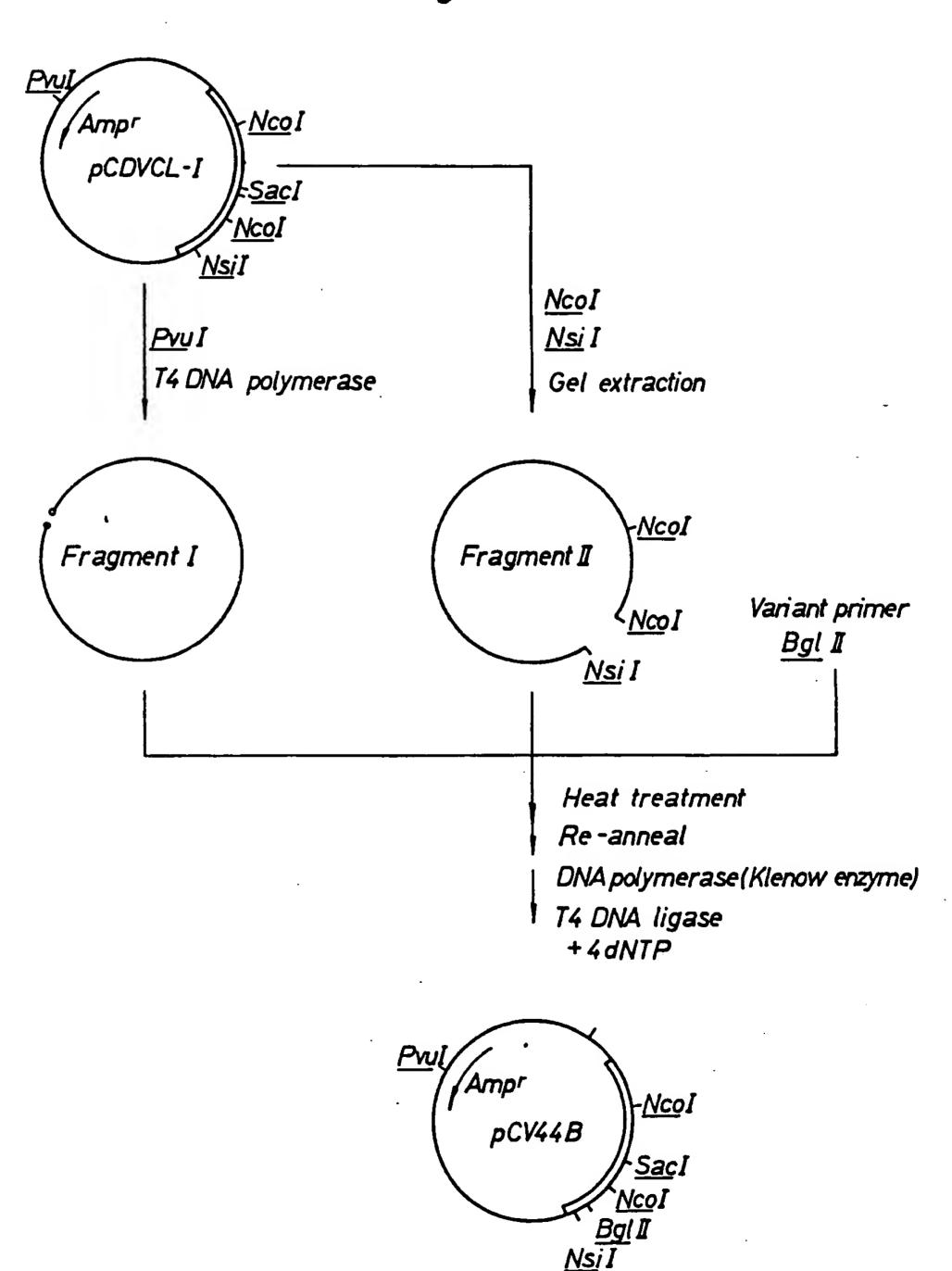
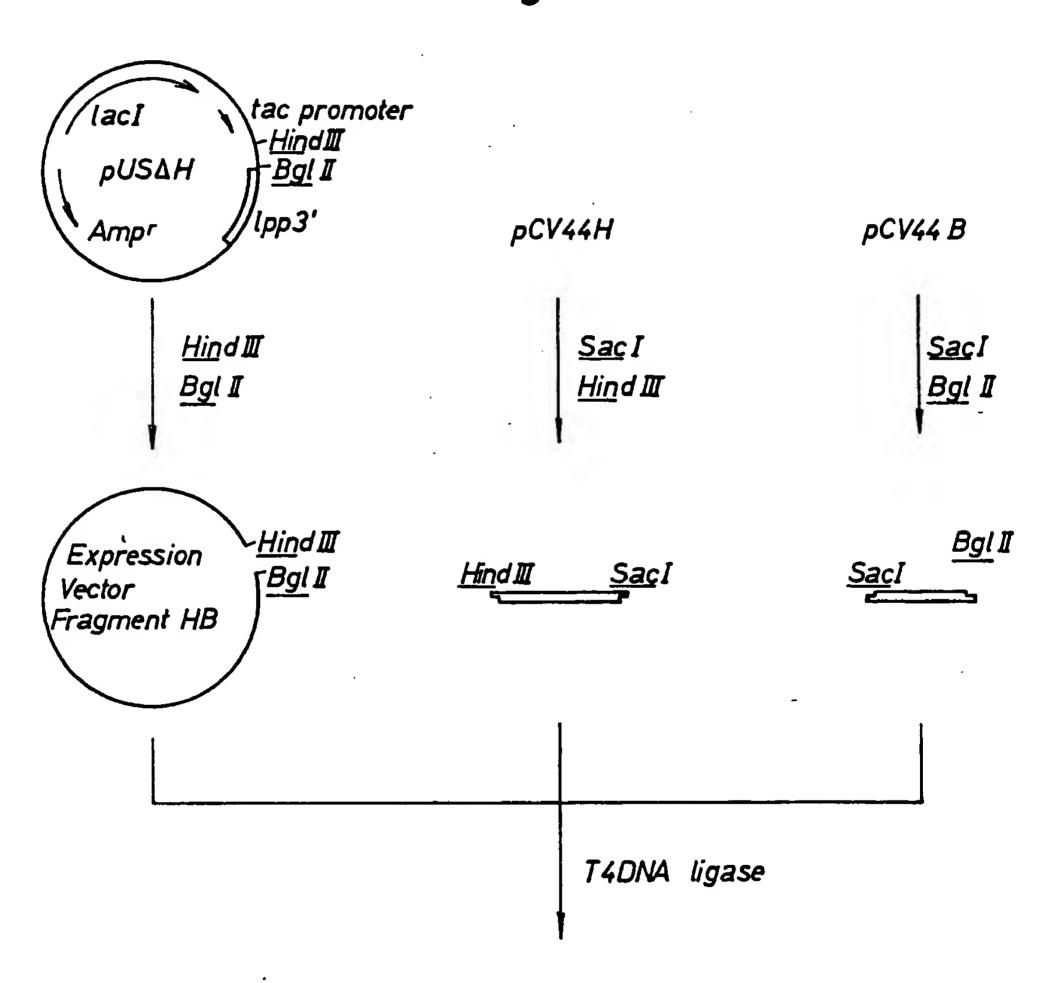
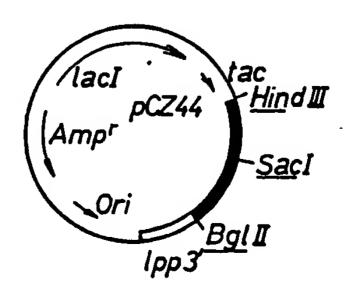


Fig.7

)







EUROPEAN SEARCH REPORT

Application Number

EP 88 40 0790

	DOCUMENTS CONS	IDERED TO BE RELEV	ANT	
Category		indication, where appropriate,	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int. Cl.4)
D,Y	EP-A-0 190 972 (M * Whole document *		1-11	C 12 N 15/00 A 61 K 39/29
D,Y	PROC. NATL. ACAD. March 1983, pages et al.: "Efficient by using antibody * Whole article *	1194-1198; R.A. YOUNG isolation of genes	1-11	7. GI N 33, 23
•	October 1986, page Columbus, Ohio, US; "Non-A, non-B hepa Ag is a normal cel	, vol. 105, no. 17, 527, no. 151110y, T. AKATSUKA et al.: titis related AN6520 lular protein mainly II" & J. MED. VIROL.	1	
	BIOLOGICAL ABSTRACT no. 4871; JI. TO "AN6520 antigen: and from liver with non & J. MED. VIROL 15(* Whole abstract *	HMATSU et al.: n antigen purified n-A. non-B hepatitis"	1	TECHNICAL FIELDS SEARCHED (Int. CL4) C 12 N
1	EP-A-0 066 296 (E)	ISAI CO., LTD)		A 61 K G 01 N
A	EP-A-0 092 249 (E)	ISAI CO., LTD)		
			·	•
	The present search report has b			
THE	HAGUE	Date of completion of the search 21-06-1988	SKELI	Examiner LY J.M.
X : partic Y : partic docus A : techn O : non-t	ATEGORY OF CITED DOCUME cularly relevant if taken alone cularly relevant if combined with an ment of the same category cological background written disclosure mediate document	E : earlier patent after the filin other D : document cit L : document cit	sciple underlying the i	avention hed on, or

O PORM 1503 00.82





(1) Publication number: 0 293 274 B1

(12)

EUROPEAN PATENT SPECIFICATION

Date of publication of patent specification: 04.09.91 Bulletin 91/36

(61) Int. Cl.5: C12N 15/12, C12N 15/33

(21) Application number: 88400790.7

2 Date of filing: 31.03.88

- 64) DNA molecules encoding non-A, non-B hepatitis antigens, and their use in producing said antigens.
- 30 Priority: 31.03.87 JP 78313/87 04.06.87 JP 140586/87 10.11.87 JP 283990/87
- Date of publication of application: 30.11.88 Bulletin 88/48
- 48 Publication of the grant of the patent: 04.09.91 Builetin 91/36
- Designated Contracting States:
 CH DE FR GB IT LI NL SE
- 68 References cited: EP-A- 0 066 296 EP-A- 0 092 249 EP-A- 0 190 972 PROC. NATL. ACAD. SCI. USA, vol. 80, March 1983, pages 1194-1198; R.A. YOUNG et al.: "Efficient isolation of genes by using antibody probes" CHEMICAL ABSTRACTS, vol. 105, no. 17, October 1986, page 527, no. 151110y, Columbus, Ohio, US; T. AKATSUKA et al.: "Non-A. non-B hepatitis related AN6520 Ag is a normal cellular protein mainly expressed in liver II * & J. MED. VIROL 1986, 20(1), 43-56 BIOLOGICAL ABSTRACTS, vol. 80, 1985, no. 4871; J.-I. TOHMATSU et al.: "AN6520 antigen: an antigen purified from liver with non-A. non-B hepatitis* & J. MED. VIROL 15(4): 357-372, 1985.

- 73 Proprietor: MITSUBISHI KASEI CORPORATION 5-2, Marunouchi 2-chome Chiyoda-ku Tokyo 100 (JP)
- (72) Inventor: Takahashi, Kazuhiro 303 Corpo **Omodaka** 29 Ichijoji-Somedono-cho Sakyo-ku Kyoto City Kyoto Prefecture (JP) Inventor: Shibui, Tatsurou M606-11 San White 6-15-4 Minami-Naruse Machida City Tokyo (JP) Inventor: Kamizono, Michiru F202-24 San White 2-8-11 Tsuruma Machida City Tokyo (JP) Inventor: Matsul. Rie A211-23 San White 1-26-31 Ogawa Machida City Tokyo (JP) Inventor: Teranishi, Yutaka 1-14-2 Takane Sagamihara City Kanagawa Prefecture (JP) Inventor: Nakanishi, Shigetada 517-116 Iwakura-Nagatani-cho Sakyo-ku **Kyoto City Kyoto Prefecture (JP)** Inventor: Kitamura, Naomi 2-272 Yakumo-Higashi-machi Moriguchi City Osaka Prefecture (JP)
- Representative: Gutmann, Ernest et al S.C. Ernest Gutmann Yves Plasseraud 67, boulevard Haussmann F-75008 Paris (FR)

EP 0 293 274 B1

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid (Art. 99(1) European patent convention).

EP 0 293 274 B1

Description

10

15

50

55

The present invention generally relates to the production of an antigen specific to non-A non-B hepatitis by recombinant DNA technology. More particularly, it relates to a DNA fragment coding for an antigen specifically occurring in a host affected with non-A non-B hepatitis, an expression vector containing such a DNA fragment, a host transformed with such an expression vector, as well as a process for producing said antigen specific to non-A non-B hepatitis by culturing such a transformant.

Among viral hepatitises, the viral entities of hepatitis type A and type B have been found and, accordingly, it has now become possible to diagnose such diseases by immunological methods.

Still another type of hepatitis different from the types A and B, which is called non-A non-B type hepatitis, is said to be over 90% of post-transfusion hepatitis: refer to NIPPON RINSHO (Japan Clinic), 35, 2724 (1977); J. Biol. Med., 49, 243 (1976). The pathogenic virus of the non-A non-B type hepatitis, however, has not yet been identified. Only one fact which has already been established is potential infection of human hepatitis type non-A non-B virus to chimpanzee: refer to Lancet I, 459 (1978); ibid., 463 (1978).

Many workers have done various investigations for searching an antigen-antibody system related to the non-A non-B hepatitis by using mainly sera from patients affected with the disease; nevertheless, no definite system has been found. Under these circumstances, the diagnosis of non-A non-B hepatitis should inevitably be effected by so-called exclusion diagnosis: that is, whether or not the hepatitis of a patient is type A or type B or other hepatitis due to a virus known to cause hepatopathy, for example, CMV, HSV, EBV, etc., is first determined; and if not, the patient's hepatitis is diagnosed as non-A non-B type. Thus, such a diagnosis of non-A non-B hepatitis will require much time and labor.

An antigenic protein specific to non-A non-B hepatitis and useful for the direct diagnoses of the hepatitis has been purified from human and chimpanzee hepatocytes affected with non-A non-B hepatitis, and a monoclonal antibody specific to the antigen and useful for the treatment of the non-A non-B hepatitis has also been proposed: refer to Japanese Patent Application Laying-open (KOKAI) Nos. 176856/86 and 56196/86.

A large amount of such an antigenic protein specific to non-A non-B hepatitis should be required when such a protein is to be employed, for example, as a diagnostic agent. However, it is not always appropriate to purify such a large amount of the antigenic protein from chimpanzee hepatocytes affected with non-A non-B hepatitis.

On the other hand, in order to detect a gene coding for a specific antigen of non-A non-B hepatitis by nucleic acid hybridization and, further, to produce such an antigen specific to non-A non-B hepatitis by the recombinant DNA technology, it is essential to obtain a gene fragment coding for the antigenic protein specific to the non-A non-B hepatitis.

The present inventors have made great efforts to produce such a specific antigenic protein in a large amount by genetic engineering techniques, and finally isolated a gene fragment coding for the antigenic protein specific to non-A non-B hepatitis, said gene fragment being useful for the production of such antigens. Further, the inventors have successfully constructed an expression vector containing said gene fragment. Thus, the present invention has now been attained.

It is an object of the invention to provide a DNA fragment which contains a base sequence coding for an antigen specifically occurring in a host cell affected with non-A non-B hepatitis or an antigenic protein specific to non-A non-B hepatitis having physiological activities equivalent to those of said specifically occurring antigen.

Another object of the invention is to provide an expression vector having said DNA fragment introduced thereinto at a cloning site downstream from a promoter of the vector.

A still another object of the invention is to provide a transformant obtained by transforming a host cell with said expression vector.

A further object of the invention is to provide a process for producing such an antigen specific to non-A non-B hepatitis by culturing said transformant.

Other objects and advantages of the present invention will be apparent from the following detailed description with reference to the attached drawings, in which:

Figs. 1a-1e show the base sequence coding for an antigenic protein specific to non-A non-B hepatitis; Fig. 2 shows the base sequence of a hybrid promoter Pac;

Figs. 3a-3c show the base sequence of a cDNA fragment obtained in Example 1 described hereinbelow, together with deduced amino acid sequence;

Figs. 4a-4c show the base sequence of cDNA containing the full length gene sequence of an antigenic protein specific to non-A non-B hepatitis, which cDNA was obtained in Example 2 described hereinbelow, the base sequence 57-1388 thereof coding for the antigenic protein specific to non-A non-B hepatitis;

Fig. 5 schematically illustrates the construction of a plasmid pCV44H:

Fig. 6 schematically illustrates the construction of a plasmid pCV44B; and

Fig. 7 schematically illustrates the construction of a plasmid pCZ44.

EP 0 293 274 B1

The present invention will be described in detail hereinbelow.

5

15

45

*5*0

According to one aspect of the invention, a DNA fragment is provided which contains a base sequence coding for an antigenic protein occurring specifically in hepatocytes affected with non-A non-B hepatitis.

Such a DNA fragment of the invention may be prepared in the following manner.

First, a liver tissue specimen derived from a human or chimpanzee individual affected with non-A non-B hepatitis is homogenized in an aqueous solution of guanidinium thiocyanate and then subjected to cesium chloride equilibrium density gradient centrifugation according to Chirgwin et al. method (Biochemistry, 18, 5294-5299 (1979)) to separate total RNA as a precipitate. After separation, the total RNA is purified by phenol extraction and ethanol precipitation.

"Individuals affected with non-A non-B hepatitis" used as sources of liver tissue specimens in the invention may include those affected with so-called type D hepatitis, which has recently been named.

It is known that mRNA of an antigen gene generally has a poly-A chain. Thus, the total RNA is subjected to oligo(dT) cellulose column chromatography in a conventional manner and poly(A)-containing RNA (poly A+RNA) is isolated as mRNA material.

A cDNA library corresponding to the poly A+ RNA is then obtained from the mRNA material according to the random primer method (Y. Ebina et al., Cell, 40, 747-758 (1980)): Thus, a number of DNAs complementary to the mRNA material are randomly synthesized using any primer of e.g. about 6 bases and a reverse transcriptase.

The cDNA is methylated with a DNA methylase, e.g. EcoRI methylase, to protect cleavage sites present in the cDNA capable of being cleaved by a corresponding restriction enzyme, e.g. EcoRI. A DNA linker containing the corresponding restriction enzyme cleavage sites at both ends, e.g. EcoRI linker (CGAATTCG), is added to the methylated cDNA and, then, this cDNA is digested with the restriction enzyme, e.g. EcoRI.

The digested cDNA is then cloned into a cloning vector such as a plasmid or a λ phage. For example, the cDNA may be introduced into EcoRI site of λgt 11 DNA, which is an expression cloning vector: refer to R.A. Young et al., Pro. Nati. Acad. Sci. U.S.A., 80, 1194-1198 (1983). The cDNA will be inserted into the β-gal gene on the λgt 11 phage. Thus, expression of the cDNA can be easily verified by the production of a fused protein with β-galactosidase due to induction of the expression by the lactose operon promoter of said phage when E. coli transfected with said phage is cultured in a medium containing IPTG (isopropylthio-β-D-galactopyranoside).

The λ gt 11 phage incorporating the cDNA is then introduced into <u>E. coli</u> by Tomizawa <u>et al.</u> method in "Experimental Procedures for Bacteriophages", pp. 99-174, published May 30, 1970 by Iwanami Shoten (Japan). The thus transfected microorganism is cultured in an IPTG-containing medium.

The thus formed piaques can be easily selected by an immunological screening method using a monoclonal antibody specifically directed to non-A non-B hepatitis to obtain a desired cDNA. Such a monoclonal antibody which can be used in the immunological screening method may be prepared according to the methods described in Japanese Patent Application Laying-open Nos. 56196/86 and 91328/88. The screening methods used may include the western blotting technique described in these applications.

The plaques positive in the immunological screening test are selected to proliferate the phage by Tomizawa et al. method. DNA is purified from the grown phage by T. Maniatis et al. method in "Molecular Cloning", Cold Spring Harbor Laboratory, pp. 85 et seq. (1982), and cleaved with a suitable restriction enzyme such as EcoRi. The thus purified and digested DNA fragments can be used to determine the base sequence of a desired cDNA segment according to Maxam and Gilbert method in Methods in Enzymology, 65, 499-560 (1980); or alternatively, after further cloning the DNA fragments into M13 phage, the base sequence of such a desired cDNA segment can be determined according to the dideoxy method: Sanger et al., Proc. Natl. Acad. Sci. U.S.A., 74, 5463 (1977).

Thus, a cDNA fragment coding for an antigen specific to non-A non-B hepatitis can be obtained. However, such a DNA fragment may usually be only a portion of the gene coding for the non-A non-B hepatitis-specific antigen.

A full length cDNA coding for such a non-A non-B hepatitis-specific antigen may be obtained in the following manner.

Poly A*-mRNA is isolated and purified in a manner similar to that described above. From the poly A*-mRNA a cDNA library is obtained according to Okayama-Berg vector-primer method: Molecular and Cellular Biology, 2, 161-170 (1982).

A plasmid containing such a cDNA thus prepared is used to transform E. coli by any conventional method, for instance, the method D. Hanahan: J. Mol. Biol., 168, 557 (1983). The transformant ampicilin-resistant strains are collected and screened by the colony hybridization method using the aforementioned DNA fragment as a probe. Such a probe may preferably be prepared by either the strepto-avidin method, or the nick translation method using photobiotinnucleic acids and ²²P-nucleic acids.

The thus selected colonies containing a cDNA clone are cultured. Plasmid DNA is obtained from the cul-

tured colony according to Birmboim et al. method (Nucleic Acid Res., 7, 1513 (1979)) and digested with a suitable restriction enzyme. The base sequence of a desired full-length cDNA segment is then determined according to the aforementioned Maxam and Gilbert method or, alternatively, after further cloning the digested DNA into M13 phage or pVC12 plasmid, such a base sequence is determined according to the above described Sanger et al. dideoxy method.

The base sequence of the full length DNA coding for an antigen specific to non-A non-B hepatitis is shown in Fig. 1, in which the symbol "—" just under the base sequence represents a corresponding base complementary to the respective base described just above each of the symbols.

Of course, DNA fragments which can be employed in the invention do not necessarily contain the same base sequence as shown in Fig. 1, but those DNA fragments in which a part of said base sequence shown in Fig. 1 has been substituted by at least one different base or deleted therefrom and those DNA fragments in which one or more additional bases have been added to the base sequence of Fig. 1 may also be included herein provided that such different DNA fragments may code for substances having physiological activities equivalent to those of the non-A non-B hepatitis-specific antigens encoded by the base sequence of Fig. 1.

According to another aspect of the invention, an expression vector is provided in which the aforementioned DNA fragment of the invention is inserted into a cloning site downstream from a promoter of this vector.

15

30

55

The expression vector of the invention contains a promoter in a position capable of controlling the transcription of a DNA fragment coding for a non-A non-B hepatitis-specific antigen obtained by the aforementioned method. The promoters used in the invention may be any promoter capable of expressing the DNA fragment in a host, and preferably of controlling the transcription of the fragment.

When a host used is a microorganism such as Escherichia coli, Bacillus subtitis, etc., the expression vector of the invention may preferably comprise a promoter, a ribosome binding sequence, a gene for a non-A non-B hepatitis-specific antigen, a transcription termination factor, and a gene controlling the promoter.

The promoter used may include those derived from E. coli, phage, etc., for example, tryptophan synthase operon (trp), lactose operon (lac), lipoprotein (lpp), recA, lambda phage P_L, P_R, T5 early gene P₂₅, P₂₆ promoter, which may also be prepared by chemical synthesis. Also included herein are hybrid promoters such as tac (trp: lac), trc (trp: lac) and Pac (phage: E. coli) shown in Fig. 2.

The ribosome binding sequence may be derived from E. coli, phage, etc., but preferably may be those synthetically prepared, for example, those containing a consensus sequence such as

AGGAGGTTTAA. SD sequence

The gene for a non-A non-B hepatitis-specific antigen may be directly employed without any modification. Preferably, an unnecessary base sequence (non-coding region) may be deleted by site-directed mutagenesis: BIO TECHNOLOGY, July, 636-639 (1984).

A transcription termination factor may not always be required in the expression vector of the invention. Preferably, the instant vector may contain a p-independent terminator, for example, ipp terminator, trp operon terminator, ribosomal RNA gene terminator, etc.

The expression vector may be derived from any conventional plasmid. Preferably, it may be derived from such a plasmid as replicating itself in E. coli or Bacillus subtilis, for example, pBR322- or pUB110-derived plasmid.

Desirably, these factors required for expression are arranged in the expression plasmid in the order of the promoter, the SD sequence, the structural gene of a non-A non-B hepatitis-specific antigen, and the transcription termination factor from 5' to 3'. A repressor gene required to control the transcription, a marker gene such as drug-resistant gene, and a plasmid replication origin may be arranged in any order in the expression vector.

The expression vector of the invention may be introduced into a host by any conventional method for transformation of E. coli, e.g., one described in Molecular Cloning, 250-253 (1982), or of Bacilius subtilis, e.g., one described in Molec. Gen. Genet., 168, 111-115 (1979) or Proc. Nat. Acad. Sci. U.S.A., 44, 1072-1078 (1958).

The resulting transformant may be cultured in any conventional medium, e.g. one described in Molecular Cloning, 68-73, (1972), at a temperature in the range of 28 to 42°C in both cases of E. coli and Bacillus subtilis. Preferably, it may be cultured at a temperature in the range of 28 to 30°C where no expression of heat shock proteins may be induced.

The desired protein thus produced may be easily purified from the host in conventional procedures. For example, the host cell may be crushed by lysozyme-surfactant or ultra-sonication, and the insoluble fractions which contain the desired non-A non-B hepatitis-specific antigen may be then collected by centrifugation, solubilized in a surfactant such as 0.01% SDS, and subjected to column chromatography using a monocional anti-

body (Japanese Patent Application Laying-open (KOKAI) Nos. 56196/86 and 176856/86.

10

40

When an eukaryotic cell such as an animal cell is employed as a host, the expression vector of the invention is preferably as follows:

The promoters used in the vector of the invention for the expression in eukaryotic cells may herein include SV40 early and late promoters; promoters of apolipoprotein E and A-I genes; promoter of heat shock protein gene (Proc. Natl. Acad. Sci. U.S.A., 78, 7038-7042 (1981)); promoter of metallothionein gene (Proc. Natl. Acad. Sci. U.S.A., 77, 6511-6515 (1980)); HSV TK promoter; adenovirus promoter, such as Ad2 major late promoter (Ad2 MLP); LTR (long terminal repeat) of retrovirus; etc. SV40 promoter and promoter of metallothionein gene are preferred.

The expression vector of the invention may contain a splice sequence comprising 5' splice junction donor site, an intron and 3' splice junction acceptor site. A common base sequence is found at all the splice junction sites (exonintron junction sites); so-called GT/AG rule that any intron region always starts from two bases GT at the donor site and terminates at two bases AG of the acceptor site has been established.

The expression vector of the invention may contain one or more splice sequences as mentioned just above. Such splice sequences may be positioned upstream or downstream of the structural gene for a non-A non-B hepatitis-specific antigen.

lilustrative examples of such splice sequences may include those DNA sequences found in exons 2 and 3 of rabbit β-globin gene (Science, 26, 339 (1979)) and mouse methallothionein-i gene containing the promoter, exons 1, 2 and 3 and introns A and B of methallothionein gene (Proc. Natl. Acad. Sci. U.S.A., 77, 6513 (1980)). The 5' and 3' splice sites may be derived from the same or different gene; for example, a sequence in which 5' splice site contained in adenovirus DNA is linked to 3' splice site derived from the gene of Ig variable region can be employed.

The expression vector of the invention also contains a polyadenylation site downstream from the structural gene of a non-A non-B hepatitis-specific antigen. Illustrative examples of the polyadenylation sites may include those derived from SV40 DNA, β-globin gene or methallothionein gene. A combined site of the polyadenylation sites of β-globin gene and SV40 DNA may be employed in the invention.

The expression vector of the invention may also contain a dominant selective marker permitting the selection of transformants. Selective markers which can be used herein may include DHFR gene imparting MTX (methotrexate) resistance to a host; tk gene of herpes simplex virus (HSV) which permits selection of tkr strains transformed therewith in HAT medium; the gene for aminoglycoside 3'-phosphotransferase from <u>E. coli</u> transposon Tn5, which imparts to a host the resistance against 3'-deoxystreptamine antibiotic G418; bovine papilloma virus gene permitting morphological discrimination by piled up growth; and aprt gene.

Alternatively, animal cells transformed with the expression vector of the invention may be selected by the cotransformation even though no selective marker is present in the vector. For this purpose, an animal cell is cotransformed with both the expression vector and a plasmid or other DNA containing a gene for such a selective marker and selected by a phenotypic trait of the gene.

Advantageously, the expression vectors may also contain a plasmid fragment having an origin of replication derived from a bacterium such as E. coli, since such vectors can be cloned in bacteria. Such plasmids may include pBR322, pBR327, pML, etc.

tion may include pKCR (Proc. Natl. Acad. Sci. U.S.A., 78, 1528 (1981)), which contains SV40 early promoter, the spiles sequence and polyadenylation site derived from rabbit β-globin gene, the polyadenylation site from SV40 early region, and the origin of replication and ampicillin resistant gene from pBR322; pKCR H2 (Nature, 307, 605 (1984)), in which the pBR322 portion of pKCR has been substituted by pBR327 fragment and the EcoRI site present in the exon 3 of rabbit β-globin gene has been converted into Hindlil site; and pBPV MT1 containing BPV gene and methallothionein gene (Proc. Natl. Acad. Sci. U.S.A., 80, 398 (1983)).

Animal cells transformed with the expression vector of the invention may include CHO cells, COS cells, and mouse L cells, C127 cells and FM3A cells.

The introduction of the expression vector of the invention into an animal cell may be carried out by transfection, microinjection, etc. Most often, the transfection may employ CaPO₄: Virology, 52, 456-487 (1973).

Animal cells transformed by introducing the expression vector of the invention may be cultured in a suspension or solid medium by conventional methods. The culture medium used is most often MEM, RPMI1640, etc.

Proteins produced in the transformed animal cells can be separated and purified in the aimost same manner as in the case of microorganisms aforementioned.

As stated, the Invention provides a transformant cell obtained by introducing the expression vector of the invention into a host cell.

Also provided according to the Invention is a process for producing a non-A non-B hepatitis-specific antigen

comprising culturing said transformant and collecting the produced and accumulated antigen.

As stated previously, a large amount of an antigenic protein specific to non-A non-B hepatitis is required when such a protein is to be utilized as a direct diagnostic agent. According to the present invention, such an antigenic protein can be produced with a low cost and a large scale without use of infected chimpanzee hepatocytes. Prior to the present invention, it has been difficult obtain such a large amount of a non-A non-B hepatitis-specific antigenic protein from hepatocytes of chimpanzees affected with non-A non-B hepatitis.

Further, the DNA fragment coding for an antigenic protein of non-A non-B hepatitis virus according to the present invention will be useful as a probe for detecting the gene of said antigenic protein by nucleic acid hybridization.

÷

EXAMPLES

10

15

20

30

The following examples will be given by way of illustration but these examples in no way limit the scope of the invention without departing the concept thereof.

EXAMPLE 1 : Preparation of cDNA Fragment Coding for Antigenic Protein Specific to Non-A Non-B Hepatitis

Poly(A)-containing RNA was prepared from chimpanzee liver according to the guanidine thiocyanate-li-thium chloride method: Cathala et al., DNA, 2, 329 (1983).

The infected liver (5 g) was taken out from a chimpanzee affected with non-A non-B hepatitis and immediately frozen by liquid nitrogen. The frozen liver was added into a Waring blender together with liquid nitrogen and ground at 3,000 rpm for 2 minutes. The ground liver specimen was further ground by a Teflon homogenizer at 5 rpm in 100 ml of a solution : 5 M guanidine thiocyanate, 10 mM EDTA, 50 mM Tris-HCl (pH 7), 8% (v/v) β-mercaptoethanoi. The thus solubilized material (20 ml) was slowly placed on 5.7 M CsCl solution (10 ml) contained in a centrifuge tube and centrifuged at 27,000 rpm for 20 hours in Hitachi RPS 28-2 rotor. The thus precipitated RNA was collected and dissolved in 10 ml of a solution : 0.1% sodium laury/sulfate, 1 mM EDTA, 10 mM Tris-HCl (pH 7.5). The RNA was extracted with phenoi-chloroform and recovered by ethanol precipitation.

The thus obtained RNA (about 3.95 mg) was dissolved in 1 ml of a solution: 10 mM Tris-HCl (pH 8.0), 1 mM EDTA. The solution was incubated at 65°C for 5 minutes, and 5 M NaCl (0.1 ml) was added. The resulting mbdure was subjected to chromatography on an oligo(dT) cellulose column (column volume of 0.5 ml, P-L Biochemical). The thus adsorbed poly(A)-containing mRNA was eluted with a solution: 10 mM Tris-HCl (pH 7.5), 1 mM EDTA. There was obtained about 100 µg of poly(A)-containing mRNA.

The thus obtained poly(A)* mRNA (10 μg) was dissolved in 50 μl of RT buffer: 20 mM Tris-HCi (pH 8.8), 0.1 M KCl, 12 mM MgCl₂, 2 mM MnCl₂. To this solution, there was added 8 μg of random primer d(N)₆ (P-L Biochemical). The resulting mixture was heated at 95°C for 3 minutes to denature the materials, which was then cooled gradually to room temperature to anneal the random primer with the mRNA. To the annealed mixture, there were aded 10 mM 4NTP (10 μl) and reverse transcriptase (225 units) from TAKARA SHUZO (Japan), and then water was added so as to make the total volume of the mixture to 100 μl. Reaction was allowed to proceed at 42°C for one hour.

To the reaction mixture (50 μl), there were added 10 mM NAD (2 μl), 10 mM 4dNTP (10 μl), RNase H (5 units), E. coli ligase (1 unit), E. coli DNA polymerase I (6.3 units), and 10x T4 DNA ligase buffer (10 μl; 0.1 M Tris-HCl, pH 7.5, 0.1 M DTT, 60 mM MgCl₂) to make the total volume to 100 μl. The mixture was allowed to react at 37°C for one hour to synthesize a double stranded DNA.

The thus obtained double stranded DNA was extracted with an equal volume of water-esturated phenol. Phenol in the aqueous layer was removed with the aid of ether followed by ethanol precipitation. The precipitate thus obtained was dissolved in 50 µl of water, and 10x T4 DNA polymerase buffer (10 µl; 0.33 M Tris-acetic acid, pH 7.9, 0.68 M potassium acetate, 0.1 M magnesium acetate, 5 mM DTT), 10 mM 4dNTP (10 µl), and T4 DNA polymerase (6 units) were added to make the total volume to 100 µl. The mixture was reacted at 37°C for one hour. There was obtained a double stranded DNA having blunt ends, which was then extracted with phenol to remove proteins and purified by ethanol precipitation as described above. The thus purified DNA was then air dried.

To the purified DNA, there were added 50 mM Tris-HCi (pH 7.5), 1 mM Na₂EDTA, 5 mM DTT (20 μ), 100 μ M S-adenosyi-L-methionine (2 μ), and 1.8 mg/ml EcoRI methylase (0.2 μ). Reaction was effected at 37°C for 15 minutes, whereby methylating the EcoRI restriction enzyme cleavage site on the DNA fragment. The reaction mixture was then heated at 70°C for 15 minutes to deactivate the enzyme.

To the reaction mixture, there was added 3'-phosphorylated EcoRi linker (GGAATTCC) in an amount of

100 molecules thereof per molecule of the synthetic DNA. There were further added 10x T4 DNA ligase buffer (5 μl; 0.5 M Tris-HCl, pH 7.5, 60 mM MgCl₂, 10 mM DTT), 0.1 M ATP (5 μl), and T4 DNA ligase (5 units) to make the total volume to 50 μl. The resulting reaction mixture was reacted at 4°C for 16 hours followed by heating at 70°C for 10 minutes to deactivate the enzyme. Then, 10x EcoRl buffer (10 μl; 15 M Tris-HCl, pH 7.5, 0.5 M NaCl, 60 mM MgCl₂), and EcoRl (100 units) were added to make the total volume to 100 μl, and the reaction mixture was reacted at 37°C for 2 hours to cut the linker. The reaction mixture was passed through Blo Gel A-50 (0.2 cm × 32 cm, Bio RAD). Elution was effected by a buffer: 10 mM Tris-HCl (pH, 7.5), 6 mM MgCl₂. Excess EcoRl linker was removed and, thus, a double stranded cDNA having EcoRl sites at both ends thereof was purified.

To the thus obtained double stranded cDNA fragment having EcoRI sites at both ends, there were added gt 11 DNA (10 μg) cleaved with EcoRI, 10x T4 DNA ligase buffer (10 μl) as described above, 0.1 M ATP (10 μl), and T4 DNA ligase (10 units) to make the total volume to 100 μl. The mixture was reacted at 4°C for 16 hours. Thus, said double stranded cDNA fragment was inserted into λgt 11 DNA.

The λ phage packaging kit (PROMEGA, Biotech) was used to introduce said DNA into λ phage particle. The procedures for packaging were effected according to the instructions of the kit.

The λ gt 11 phage having said DNA packaged thereinto was used to transfect E. coli strain Y1090 to form plaques according to the conventional Tomizawa et al. methods described in "Experimental Procedures for Bacteriophages", pp. 99-174, published May 30, 1970 by Iwanami Shoten (Japan). Among about 200,000 plaques, one positive clone was selected by immunological screening as described hereinbelow. A monoclonal antibody used in the immunological screening was prepared by the method described in Japanese Patent Application Laying-open (KOKAI) No. 91328/88.

E. coli Y1090 (R.A. Young et al., Pro. Nati. Acad. Sci. U.S.A., 80, 1194-1198 (1983), which had been transfected with Agt 11, was inoculated in a petri dish together with soft agar held at 42°C. The transfected cell was allowed to stand at 42°C for 5 hours. A nitrocellulose filter (S & S, BA-83, pore size of 0.2 µm) containing 10 mM IPTG was placed on the cell in the dish and incubation was effected at 37°C for 3-4 hours. This nitrocellulose filter was lightly rinsed with TBS buffer (10 mM Tris-HCl, pH 7.5, 50 mM NaCl), immersed in the TBS buffer (400 ml) containing 3% gelatine and shaked at 40°C for one hour. Thus, the nitrocellulose filter was blocked. Then, a monoclonal antibody ($OD_{280} = 4.3$) directed to a non-A non-B hepatitis-specific antigen was added to TBS buffer containing 1% gelatine with a dilution of 1/400. This mixture was put into a vinyl bag together with the filter in a proportion of 2 mi of the mixture per filter, and reaction was allowed to proceed at room temperature for 16 hours. The reaction mixture was three times washed with TBS buffer (400 ml) containing 0.05% Tween 20 over 10 minutes. A labelled secondary antibody, anti-mouse IgG-PAP (horseradish peroxidase, Bio Rad) was added to TBS buffer containing 1% gelatine with a dilution of 1/1,000. This mixture and the filter were put into a vinyl bag with a proportion of 2 ml of the mixture per filter. Reaction was allowed to proceed at room temperature for 2 hours. The reaction mixture was three times washed with TBS buffer (400 ml) containing 0.05% Tween 20 over 10 minutes, in the same manner as described above. Color development was effected by dipping the filter and 4-chloro-1-naphthol (12 mg, Bio Rad) into 20 ml of TBS buffer containing hydrogen peroxide. After completion of the color development, the filter was thoroughly washed with water and put into a vinyl bag containing water. The bag was stored in a dark and cold place.

Thus, one positive plaque was obtained. The plaque was three times subjected to single plaque isolation. In each time, immunological screening was effected in the same manner as described above, verifying that the plaque was in fact positive.

40

The phage was then cultured in a large scale to purify the DNA in the following manner: First, E. coli Y1090 was cultured overnight in 10 ml of NZ medium prepared by adding NZ amine (10 g), NaCl (5 g) and 5 mM MgCl₂ to one liter of water followed by adjusting the pH to 7.2. The culture (1 ml) was transfected with the phage, with the m.o.i. (multiplicity of infection) being 0.1. The transfected culture was allowed to stand at 37°C for 10 minutes and then transferred to one liter of NZ medium. Shaking culture was effected at 37°C for 7-8 hours until the cells were lysed. Chloroform (5 ml) was added to the culture and shaking was continued for additional 30 minutes. The culture was subjected to centrifugation at 6,500 rpm for 10 minutes to remove cell debris.

NaCl (29 g) and polyethylene glycol (70 g) were added to and thoroughly dissolved in the obtained supernatant, and the solution was allowed to stand at 4°C overnight. The precipitate was collected by centrifugation at 6,500 rpm for 20 minutes, drained thoroughly, and dissolved in 20 ml of TM buffer: 10 mM Tris-HCl (pH 7.5), 5 mM MgCl₂. DNase I and RNase A were added to the solution, both with a concentration of 10 µg/ml, and the reaction was effected at 37°C for one hour. Chloroform (20 ml) was then added to the reaction mixture and stirred; thus, polyethylene glycol was distributed in the chloroform layer which was then separated from the aqueous layer. This aqueous layer was ultra-centrifuged at 28,000 rpm for 60 minutes. Thus, a pellet of phage particles was obtained.

This pellet was dissolved in TM buffer (1 ml) and subjected to CsCl density gradient centrifugation at 33,000

rpm for 20 hours. The resultant fraction containing the phage particles ($\rho = 1.45-1.50$) was dialyzed overnight against TM buffer. Proteinase K was added to the dialyzate in an amount of 100 μ g/ml and reaction was effected at 37°C for one hour. Thereafter, an equal volume of water-saturated phenol was added and phenol-extraction was gently effected. After centrifugation at 6,500 rpm for 10 minutes, the aqueous layer was removed, put into a dialysis tube, and dialyzed overnight against water at 4°C. Thus, about 5 mg of DNA was obtained.

Cleavage reaction of this DNA (100 μg) with EcoRI (100 units) in the aforementioned buffer (100 μl) at 37°C revealed that two cDNA segments of 390 bp and 345 bp were inserted into the phage DNA.

These two EcoRI fragments were re-cloned into EcoRI site of a cloning vector pUC 119. Base sequences of these DNA fragments were determined by the dideoxy method using commercially available primers CAG-GAAACAGCTATGAC and AGTCACGACGTTGTA, respectively. The base sequence of the linking portion between these two DNA fragments was similarly determined by cutting this cDNA fragment at BamHi and EcoRV sites present therein with corresponding specific restriction enzymes, inserting the resulting BamHi-EcoRV DNA fragment between BamHi and Smal sites of the plasmid pUC 119, and sequencing the fragment by the dideoxy method.

The base sequence of said cDNA fragment is shown in Fig. 3. This was a partial cDNA fragment of a gene coding for an antigenic protein specific to non-A non-B hepatitis.

EXAMPLE 2: Preparation of cDNA Containing the Full Length Gene Sequence

15

20

50

Messenger RNA was prepared as described in Example 1 and cDNA was synthesized using Okayama vector according to the conventional method described in Molecular Cloning, p. 211 et seq. The procedures used to synthesize cDNA were as follows:

To 300 µl of a solution (10 mM Tris-HCl, pH 7.5, 6 mM MgCl₂, 10 mM NaCl), there were added 400 µg of pCDV 1 (Okayama and Berg, Mol. Cell. Biol., 3, 280 (1983)) and 500 units of KpnI (TAKARA SHUZO, Japan), all restriction enzymes used hereinafter having been manufactured by TAKARA SHUZO (Japan) unless otherwise noted. Reaction was effected at 37°C for 6 hours to cut the plasmid at KpnI site therein. After phenol-chloroform extraction, ethanol precipitation was effected to recover DNA.

The DNA (about 200 µg) cleaved with KpnI was added to 200 µl of a solution which was obtained by adding dTTP in a concentration of 0.25 mM to a buffer (TdT buffer): 40 mM sodium cacodylate, 30 mM Tris-HCI (pH 6.8), 1 mM CaCl₂, 0.1 mM dithiothreitol (DTT). Further, 81 units of terminal deoxynucleotidyl transferase (TdT, manufactured by P-L Biochemicals) was also added. Reaction was effected at 37°C for 11 minutes. Thus, a poly(dT) chain (about 67 deoxythymidylic acid residues) was added to the 3' end at the KpnI-cleaved site of pCDV 1. After phenol-chloroform extraction and ethanol-precipitation, about 100 µg of pCDV 1 DNA to which poly(dT) chain had been added was recovered from the reaction mixture.

The thus obtained DNA was added to 150 µl of a buffer (10 mM Tris-HCl, pH 7.5, 6 mM MgCl₂, 100 mM NaCl), and <u>Hpal</u> (360 units) was also added, followed by reaction at 37°C for 2 hours. The reaction mixture was subjected to electrophoresis on agarose gel to separate and recover about 3.1 Kbp DNA fragment. Thus, there was obtained about 60 µg of poly(dT)-containing pCDV 1.

The thus obtained DNA was dissolved in 500 µl of a solution (10 mM Tris-HCl, pH 8.0, 1 mM EDTA), incubated at 65°C for 5 minutes, and cooled on ice. After adding 5 M NaCl (50 µl), the mixture was subjected to chromatography on oligo(dA) cellulose column (Colaborative Research). DNA having a poly(dT) chain of sufficient length was adsorbed on the column and eluted with a solution: 10 mM Tris-HCl (pH 8.0), 1 mM EDTA. Thus, there was obtained 27 µg of pCDV 1 to which poly(dT) chain had been added, abbreviated hereinafter as vector primer.

A linker DNA was prepared in the following manner: To 200 µl of a solution (10 mM Tris-HCl, pH 7.5, 6 mM MgCl₂, 50 mM NaCl), there were added about 14 µg of pL 1 (Okayama and Berg, Mol. Cell. Biol., 3, 280 (1983)) and 50 units of Pstl. Reaction was effected at 37°C for 4 hours to cut the pL 1 DNA at Pstl site. Phenol-chloroform extraction and ethanol precipitation of the reaction product gave about 13 µg of pL 1 DNA cleaved at Pstl site.

The thus obtained DNA (about 13 µg) was added to 50 µi of the TdT buffer containing dGTP at a final concentration of 0.25 mM, and 54 units of TdT (P-L Blochemicals) was also added. The mixture was incubated at 37°C for 13 minutes to add a (dG) chain (about 14 deoxyguanytic acid residues) to the 3' end at the Psti-cleaved site of pL 1. After phenol-chloroform extraction, DNA was recovered by ethanol precipitation.

The thus obtained DNA was added to 100 µl of a buffer (10 mM Tris-HCl, pH 7.5, 6 mM MgCl₂ 60 mM NaCl), and 80 units of Hindill was also added. The mixture was incubated at 37°C for 3 hours to cut the pL 1 DNA at Hindill site. The reaction product was fractionated by agarose get electrophoresis. About 0.5 Kb DNA fragment was recovered by the DEAE paper method: Dretzen et al., Anal. Biochem., 112, 295 (1981). Thus, there was obtained an oligo(dG) chain-containing linker DNA, hereinafter abbreviated simply as linker DNA.

The aforementioned poly(A)* RNA (about 2 µg) prepared in the same manner as in Example 1 and the vector primer (about 1.4 µg) were dissolved in 22.3 µl of a solution : 50 mM Tris-HCl (pH 8.3), 8 mM MgCl₂, 30 mM KCl, 0.3 mM DTT, 2 mM dNTP (dATP, dTTP, dGTP and dCTP) and 10 units of ribonuclease inhibitor (P-L Biochemicals). To the solution, there was added 10 units of reverse transcriptase manufactured by SEIKAGAKU KOGYO (Japan). Incubation was effected at 37°C for 40 minutes to synthesize a DNA complementary to the mRNA. After phenol-chloroform extraction and ethanol precipitation, the vector primer DNA to which a double stranded RNA-DNA had been added was recovered.

The thus obtained vector primer DNA containing RNA-DNA double stranded chain was dissolved in 20 μl of TdT buffer containing 60 μM dCTP and 0.2 μg poly(A). After adding 14 units of TdT (P-L Biochemical), the mbxture was incubated at 37°C for 8 hours to add a (dC) chain of 12 deoxycytidylic acid residues to the 3' end of the cDNA. The reaction product was extracted with phenoi-chloroform and precipitated with ethanol to recover a cDNA-vector primer DNA to which a (dC) chain had been added.

The thus obtained (dC) chain-containing cDNA-vector primer DNA was dissolved in 400 µl of a solution (10 mM Tris-HCl, pH 7.5, 6 mM MgCl₂, 60 mM NaCl), and 20 units of Hindlli was also added. The mixture was incubated at 37°C for 2 hours to cut the DNA at Hindlli site. The reaction product was extracted with phenol-chloroform and precipitated with ethanol. Thus, there was obtained 0.5 pmole of a (dC) chain-containing cDNA-vector primer DNA.

The thus obtained (dC) chain-containing cDNA-vector primer DNA (0.08 pmole) and the aforementioned linker DNA (0.16 pmole) were dissolved in 40 μl of a solution: 10 mM Tris-HCl (pH 7.5), 0.1 M NaCl, 1 mM EDTA. The resulting solution was incubated at 65°C for 10 minutes, at 42°C for 25 minutes, and then at 0°C for 30 minutes. The reaction mixture was adjusted to 20 mM Tris-HCl (pH 7.5), 4 mM MgCl₂, 10 mM (NH₄)₂SO₄, 0.1 M KCl and 0.1 mM β-NAD in a total volume of 400 μl.

To the reaction mixture, there was added 10 units of <u>E. coli</u> DNA ligase (New England Biolabs), followed by incubation overnight at 11°C. After adjusting the concentrations of dNTP and β-NAD in the reaction mixture to 40 μM and 0.15 mM, respectively, by supplementing necessary reagents, 5 units of <u>E. coli</u> DNA ligase, 7 units of <u>E. coli</u> DNA polymerase I (P-L Biochemicals) and 2 units of <u>E. coli</u> ribonuclease H (P-L Biochemicals) were added to the reaction mixture. The mixture was incubated at 12°C for one hour and then at 25°C for one hour.

In the course of the above reactions, a recombinant DNA containing the cDNA was cyclized and the RNA portion of the RNA-DNA double stranded chain was substituted by DNA. Thus, a desired recombinant plasmid containing a complete double-stranded DNA was produced.

The recombinant plasmid was used to transform competent cells of <u>E. coli</u> strain MC1064 prepared by conventional methods. Approximately 50,000 transformants were fixed on a nitrocellulose filter. These colonies were screened according to the colony hybridization method described in Molecular Cloning, Cold Spring Harbor Laboratory, p. 329 et seq. (1982) using the cDNA fragment obtained in Example 1 as a ³²P-labelled probe. Thus, three clones showed strong hybridization at 42°C.

These positive clones were analyzed in detail by Southern method: J. Mol. Biol., <u>98</u>, 503 (1975). There was obtained the desired full length cDNA of a gene coding for an antigenic protein specific to non-A non-B hepatitis. The base sequence of the cDNA is shown in Fig. 4.

The expression vector containing the full length cDNA was designated as pCDVCL-I.

EXAMPLE 3: Preparation of Expression Vector and Transformant and Expression of Specific Antigen

A. Preparation of Expression Vector and Transformant

1) Modification of N-terminus (Fig. 5):

40

45

50

55

i) In 100 μl of a buffer (10 mM Tris-HCI, pH 7.5, 100 mM NaCi, 6 mM MgCl₂), pCDVCL-i (5 μg) was digested with Pvui (10 units) at 37°C for 2 hours. The reaction mixture was heated at 75°C for 15 minutes to deactivate the enzyme, dialyzed against water, and dried. The cleaved plasmid DNA was treated with T4 DNA polymerase (4 units) in 40 μl of a system : 33 mM Tris-acetic acid (pH 7.9), 66 mM potassium acetate, 10 mM magnesium acetate and 0.5 mM dithiothreitol, to which 2 mM 4-deoxytriphosphate had been added; thus, the 3' protruding end of the plasmid DNA was filled in to produce a blunt end. The thus treated mixture was heated at 70°C for 10 minutes to deactivate the enzyme, dialyzed against water, and dried. The thus obtained plasmid DNA was then stored in the form of an aqueous solution (50 μl). This plasmid DNA fragment is hereinafter designated as Fragment i.

ii) On the other hand, pCDVCL-I (20 μg) was digested with Ncol and Hindlii (each 20 units) at 37°C for 2 hours in 100 μl of a buffer: 10 mM Tris-HCi (pH 7.5), 100 mM NaCl, 6 mM MgCl₂. The plasmid DNA was

subjected to 5% acrylamide gel electrophoresis at 10 V/cm for 1.5 hours in a buffer: 89 mM Tris, 89 mM boric acid, 2 mM EDTA. The gel was stained with 0.05% aqueous ethidium bromide solution and two gel slices corresponding to DNA fragments of larger molecular weights were excised from the gel under ultraviolet radiation at 340 nm. The gel slices were crushed by means of a glass rod, suspended into 4 mi of a buffer for DNA extraction (0.5 M ammonium acetate, 10 mM magnesium acetate, 1 mM EDTA, 0.1% sodium laurylsulfate), and allowed to stand overnight at 37°C to extract DNA from the gel. The materials were subjected to centrifugation at 10,000 rpm for 15 minutes to eliminate larger gel pieces, and passed through a glass filter to remove smaller gel pieces. The DNA was purified by effecting ethanol precipitation three times and stored in the form of an aqueous solution (200 µl). This plasmid DNA fragment is hereinafter designated as Fragment II.

iii) A primer of the DNA portion to be modified as shown below (51 bases) was synthesized by a DNA synthesizer, NIKKAKI (Japan), Applied Biosystem MODEL 380A. The synthesized DNA was overnight reacted with concentrated aqueous ammonia at 55°C to deprotect and purified by reversed HPLC before use.

CATGGTTGCATG wherein x represents a base substitution.

The synthetic primer (150 pmole) was treated with T4 polynucleotide kinase (20 units) in 10 μl of a kinase buffer (50 mM Tris-HCl, pH 8.0, 10 mM MgCl₂, 5 mM dithiothreitol) to phosphorylate the 5' end thereof. iv) Fragment I (0.05 pmole), Fragment II (0.05 pmole) and 5'-phosphorylated primer (45 pmole) were added to 12 μl of 5x polymerase-ligase buffer (0.5 M NaCl, 32.5 mM Tris-HCl, pH 7.5, 40 mM MgCl₂, 5 mM β-mercaptoethanol) to make the total volume of the mixture 34.8 μl. The mixture was boiled at 100°C for 3 minutes, immediately after which it was placed in a thermostat at 30°C and allowed to stand for 30 minutes. The mixture was allowed to stand at 4°C for 30 minutes and then on ice for 10 minutes to form a heteroduplex.

To an aqueous solution (11.6 μ) containing the heteroduplex, there were added 2.5 mM 4-deoxynucleotide triphosphate (2 μ I), 10 mM ATP (2 μ I), Klenow enzyme (2 units) and T4 DNA ligase (0.5 units) to form a mixture of 20 μ I in total volume. The mixture was reacted overnight at 16°C to cyclize the DNA.

An aqueous solution (2 µl) containing the circular DNA was used to transform E. coli HB101 strain according to conventional methods. Plasmids were separated from the transformant and purified in conventional manners. The plasmid was cleaved with restriction enzyme Hindlii and subjected to 5% acrylamide gel electrophoresis. Thus, two separate fragments were collected as desired modified, variant plasmids. Since resulting variant plasmids might often be admixed with original wild-type plasmids, the thus obtained variant plasmids were again employed to transform E. coli HB101 so as to purify the plasmid.

Thus, a purified plasmid pCV44H was obtained (Fig. 5).

45 II) Modification of C-terminus (Fig. 6):

i) Plasmid pCDVCL-I (5 µg) was treated in the same manner as in i) i) described above to produce Fragment

ii) Plasmid pCDVCL-I (20 μg) was treated in the same manner as in i) ii) described above except that Ncol and Nsil (each 5 units) were employed. Thus, Fragment II was produced.

iii) In the same manner as in I) iii) described above, the following primer (46 bases) was synthesized and the 5' end thereof was phosphorylated.

٠,

55

50

5

10

25

Primer

GCACAAGGAAAAAATGAGATCTGTCGACGGTTCACGTA

XXXXXXXXXXXX

(Original sequence) (-----AGATATGTGAA*A----
AATTTCC wherein x represents a base substitution and * represents substitution and * represents a base substitution and * rep

iv) The Fragment I and II and the 5'-phosphorylated primer obtained above in II) i) to iii) were treated in the same manner as in I) iv) described above. Thus, plasmid pCV44B was obtained (Fig. 6).

III) Introduction of cDNA coding for specific antigen into expression vector (Fig. 7):

i) In 100 µl of a buffer H (10 mM Tris-HCl, pH 7.5, 100 mM NaCl, 6 mM MgCl₂), 10 µg (about 3 pmole) of pCV44H was cut with Hindill (20 units) and Sacl (20 units) at 37°C for 2 hours. The reaction mixture was subjected to 5% acrylamide gel electrophoresis. Thus, a 467 bp DNA fragment coding for the N-terminus of the specific antigen was separated and purified. This fragment is hereinafter designated as Fragment N

ii) In 100 μl of the buffer H, 10 μg (about 3 pmole) of pCV44B was cleaved with <u>Bglll</u> (20 units) and <u>Sacl</u> (20 units) at 37°C for 2 hours. The reaction mbeture was subjected to 5% acrylamide gel electrophoresis to isolate and purify a 836 bp DNA fragment coding for the C-terminus of the specific antigen. The thus obtained fragment is hereinafter designated as Fragment C.

iii) In 20 μl of buffer H, 2 μg (about 1 pmole) of an expression vector pUSΔH was cut with HindIII (2 units) and BglII (2 units) at 37°C for 2 hours. The reaction mixture was extracted with an equal volume of water-saturated phenol to remove proteins. After extracting the phenol with ether, the reaction mixture was dialyzed against water to desait, and concentrated by a vacuum pump. Thus, there was obtained 10 μl of an aqueous solution containing an expression vector fragment HB.

lv) Fragment N (0.5 pmole), Fragment C (0.5 pmole) and the expression vector fragment HB (0.1 pmole) were mixed and reacted with T4 DNA ligase (1 unit) at 4°C for 16 hours in 10 μi of a buffer (10 mM Tria-HCl, pH 7.5, 1 mM dithiothreitol, 6 mM MgCl₂, 1 mM ATP). The reaction mixture (3 μl) was used to transform commercially available <u>E. coli</u> JM109 competent cell according to conventional methods. The resulting transformants were selected in L broth plate (bactopeptone 10 g, yeast exstract 5 g, NaCl 10 g, agar 15 g per liter) containing 20 μg/ml ampicillin. Thus, there was obtained an expression vector pCZ44 containing

the specific antigen gene inserted thereinto (Fig. 7).

B. Expression of Specific Antigen

20

25

30

35

E. coli strain JM109 possessing pCZ44 was cultured overnight at 30°C in L broth. The culture was inoculated in a fresh L broth with a dilution of 1/50 and cultured with shaking at 30°C for 2 hours. After IPTG (isopropylthlo-β-D-galactopyranoside) was added to the medium in a concentration of 2 mM, shaking culture was continued at 30°C for further 3 hours. The cells were collected by centrifugation at 6,500 rpm for 10 minutes and suspended in a buffer (0.9% NaCl, 10 mM Tris-HCl, pH 7.5) to store.

C. Verification of Expression of Specific Antigen

The thus obtained cell culture (0.3 ml) was subjected to 10% SDS polyacrylamide gel electrophoresis at 120 V for one hour in a buffer (Tris 3g/l, glycine 14.4 g/l, 0.1% SDS). The gel was removed, placed on a nit-rocellulose filter, interposed between filter papers and electrophoresed at 5 V/cm, 4°C in a buffer (Tris 3g/l, glycine 14.4 g/l) to transferred proteins in the gel onto the nitrocellulose filter. The nitrocellulose filter was rinsed with TBS buffer (10 mM Tris-HCl, pH 7.5, 50 mM NaCl), immersed into 400 ml of TBS buffer containing 3% gelatine and shaked at 40°C for one hour to block the nitrocellulose filter.

To TBS buffer containing 1% gelatine, there was added a monoclonal antibody directed to a non-A non-B hepatitis-specific antigen ($OD_{200} = 4.3$) with a dilution of 1/400. The resulting mixture and the nitrocellulose filter were put into a vinyl bag so that the mixture was present in an amount of 2 ml per filter. Reaction was effected at room temperature for 16 hours. The reaction mixture was washed three times with 400 ml of TBS buffer con-

taining 0.05% Tween 20 for 10 minutes.

To TBS buffer containing 1% gelatine, there was added a labelled secondary antibody, anti-mouse IgG-PAP (horseradish peroxidase, Bio Rad), with a dilution of 1/1000. The resulting mixture and the nitrocellulose filter were put into a vinyl bag so that the mixture was present in an amount of 2 ml per filter. Reaction was effected at room temperature for 2 hours. The reaction mixture was washed three times with 400 ml of TBS buffer containing 0.05% Tween 20 for 10 minutes.

Color formation was effected by immersing the filter into 20 ml of TBS buffer containing 12 mg of 4-chlo-ro-1-naphthoi (Bio Rad) and hydrogen peroxide. After completion of color formation, the filter was thoroughly washed with water, put into a vinyl bag containing water, and stored in a dark and cold place.

Such a test effected showed that a protein reacting with the monoclonal antibody was found at the same position (44 Kd) as found in the case of the specific antigen derived from infected chimpanzee liver. This verifies that such a specific antigen can be in fact expressed in E coli. The invention thus also relates to a process for the in vitro diagnosis of NON-A NON-B hepatitis, which comprises contacting a liver sample and/or a serum sample taken from a patient possibly infected with a NON-A NON-B hepatitis with the protein whose formula appears in claim 3 hereafter or a part thereof for a time and under conditions sufficient to allow for the production of a complex between said protein or part thereof with the antibodies contained in the patient sample and detecting the presence of the immunologic complex, particularly when the patient is suffering from NON-A NON-B hepatitis.

Any part of said protein, or any recombinant, produced by genetic engineering and including the aminoacid sequence of said protein or part of said protein can be substituted for above-said protein, it being understood that the said recombinant protein or part of said protein are specifically recognized by the same antibodies as those which recognize said protein.

In other words the invention relates to all recombinant proteins or protein fragments which bind to antibodies contained in a liver extract or serum sample, or both, and originating from a patient suffering from NON-A NON-B hepatitis.

The invention also relates to a process for detecting in vitro an infection by a NON-A NON-B hepatitis virus, which process comprises contacting the DNA of claim 4, or a fragment thereof, under suitable hybridization conditions, with a sample of liver extract and/or serum sample originating from the patient to be diagnosed and in which the nucleic acid components had previously been made accessible to hybridization, to form a hybridization product between said DNA of claim 4 (probe) and the viral DNA of a NON-A NON-B hepatitis B virus, and detecting said hybridization product, particularly in the case where the patient is indeed infected with a NON-A NON-B virus.

35 Claims

10

1. A DNA fragment which contains a base sequence coding for an antigenic protein specifically occurring in a host affected with non-A non-B hepatitis, said protein comprising the whole or a part of the amino acid sequence represented by the formula:

*5*5

10 Met Ala Val Thr Thr Arg Leu Thr Trp Leu His Glu Lys Ile Leu 5 20 Gln Asn His Phe Gly Gly Lys Arg Leu Ser Leu Leu Tyr Lys Gly Ser Val His Gly Phe His Asn Gly Val Leu Leu Asp Arg Cys Cys 50 Asn Gln Gly Pro Thr Leu Thr Val Ile Tyr Ser Glu Asp His Ile 15 Ile Gly Ala Tyr Ala Glu Glu Gly Tyr Gln Glu Arg Lys Tyr Ala 90 Ser Ile Ile Leu Phe Ala Leu Gln Glu Thr Lys Ile Ser Glu Trp 20 100 Lys Leu Gly Leu Tyr Thr Pro Glu Thr Leu Phe Cys Cys Asp Val 110 120 Ala Lys Tyr Asn Ser Pro Thr Asn Phe Gln Ile Asp Gly Arg Asn 130 Arg Lys Val Ile Met Asp Leu Lys Thr Met Glu Asn Leu Gly Leu . 30 150 Ala Gln Asn Cys Thr Ile Ser Ile Gln Asp Tyr Glu Val Phe Arg 160 Cys Glu Asp Ser Leu Asp Glu Arg Lys Ile Lys Gly Val Ile Glu 35 180 Lau Arg Lys Ser Leu Leu Ser Ala Leu Arg Thr Tyr Glu Pro Tyr 190 Gly Ser Leu Val Gln Gln Ile Arg Ile Leu Leu Gly Pro Ile 40

45

__

50

55

•

210 200 Gly Ala Gly Lys Ser Ser Phe Phe Asn Ser Val Arg Ser Val Phe 220 Gln Gly His Val Thr His Gln Ala Leu Val Gly Thr Asn Thr Thr 230 240 Gly Ile Ser Glu Lys Tyr Arg Thr Tyr Ser Ile Arg Asp Gly Lys Asp Gly Lys Tyr Leu Pro Phe Ile Leu Cys Asp Ser Leu Gly Leu 260 270 15 Ser Glu Lys Glu Gly Gly Leu Cys Met Asp Asp Ile Ser Tyr Ile 280 Leu Asn Gly Asn Ile Arg Asp Arg Tyr Gln Phe Asn Pro Met Glu 290 300 Ser Ile Lys Leu Asn His His Asp Tyr Ile Asp Ser Pro Ser Leu 310 Lys Asp Arg Ile His Cys Val Ala Phe Val Phe Asp Ala Ser Ser 25 320 330 Ile Glu Tyr Phe Ser Ser Gln Met Ile Val Lys Ile Lys Arg Ile 340 Arg Arg Glu Leu Val Asn Ala Gly Val Val His Val Ala Leu Leu 350 Thr His Val. Asp Ser Met Asp Leu Ile Thr Lys Gly Asp Leu Ile 370 Glu Ile Glu Arg Cys Val Pro Val Arg Ser Lys Leu Glu Glu Val 380 390 Gln Arg Lys Leu Gly Phe Ala Leu Ser Asp Ile Ser Val Val Ser 40 400 Asn Tyr Ser Ser Glu Trp Glu Leu Asp Pro Val Lys Asp Val Leu 410 45 Ile Leu Ser Ala Leu Arg Arg Met Leu Trp Ala Ala Asp Asp Phe Leu Glu Asp Leu Pro Phe Glu Gln Ile Gly Asn Leu Arg Glu Glu *5*0 Ile Ile Asn Cys Ala Gln Gly Lys Lys ***.

^{2.} The DNA fragment in accordance with claim 1, in which the base sequence comprises the whole or a part of the base sequence represented by the formula:

														4	_
	!51	ATG	GCA	GTG	ACA	ACT	CGT	TTG	ACA	TGG	TTG '	CAT	GAA	AAG	ATC
5	; 3 '														
				50			60			7	'n			80	
		CTG	CAA	_		TTT									TAT
10							_								
10						10								com	CNC
		AAG	GGT	AGT	GTC	CAT	GGA	TTC	CAT	AAT	GGA	GTT	TTG	CTT	GAC
		1:	30		1	140			150			10	60		
15		AGA	TGT	TGT	AAT	CAA	GGG	CCT	ACT	CTA	ACA	GTG	ATT	TAT	AGT
		170			180			19	90		:	200			210
		170 GAA				ATT									
20															
20				CAT	ATT	ATT	GGA	GCA	TAT	GCA	GAA	GAG	GGT	TAC	CAG
20		GAA	GAT	CAT	ATT	ATT	GGA	GCA 	TAT	GCA	GAA 240	GAG	GGT	TAC	CAG
20		GAA	GAT	CAT	ATT	ATT	GGA	GCA 	TAT	GCA	GAA 240	GAG	GGT	TAC	CAG
20		GAA	GAT	CAT	ATT	ATT	GGA	GCA 	TAT	GCA	GAA 240	GAG	GGT	TAC	CAG
		GAA	GAT AGA	CAT Z: AAG	ATT 20 TAT	ATT	GGA TCC	GCA 230 ATC	ATC	GCA CTT	GAA 240 TTT	GAG	CTT	TAC 2 CAA	CAG 50 GAG
		GAA	GAT AGA	CAT Z: AAG	ATT 20 TAT	ATT	GGA TCC	GCA 230 ATC	ATC	GCA CTT	GAA 240 TTT	GAG	CTT	TAC 2 CAA	CAG 50 GAG
		GAA	GAT AGA	CAT Z: AAG	ATT 20 TAT	ATT	GGA TCC	GCA 230 ATC	ATC	GCA CTT	GAA 240 TTT	GAG	CTT	TAC 2 CAA	CAG 50 GAG
		GAA	GAT AGA	ZAAG AAG 260 ATT	ATT 20 TAT	GCT	GGA TCC	GCA 230 ATC	ATC	GCA CTT	GAA 240 TTT	GCA	CTT	2 CAA 290 CCA	CAG 50 GAG
25		GAA	AAA 300	ZAAG AAG 260 ATT	ATT 20 TAT TCA	GCT GAA	TCC 270 TGG	GCA 230 ATC	ATC	GCA CTT 2 GGA 320	GAA 240 TTT 80 CTA	GCA	GGT CTT ACA	2 CAA 290 CCA	CAG 50 GAG

	34	10		3	150			360			37	ra		
			CAG										ATG	GAC
				~~~										
5														
	380	. •		390			40	0		4	110			420
	TTA	AAG	ACA	ATG	GAA	AAT	CTT	GGA	CTT	GCT	CAA	TAA	TGT	ACT
				***										
10							40							
10	100		43											
	ATC	TCT	ATT	CAG	GAT	TAI	UAA 	GTT	TTT	CGA	TGC	GAA	GAT	TCA
		4	70			480			49	0		5	00	
15,	CTG		GAA											AAG
												540		
20	AGC	TTA	CTG	TCT	GCC	TTG	AGA	ACT	TAT	GAA	CCA	TAT	GGA	TCC
	55	50		•	60			570			86	20		
		GTT	CAA	CAA	ATA	CGA	ATT	CTG	CTG	CTG	CCT	CCA	<b>ATT</b>	GGA
25														
	590			600			61	LO			520			630
	GCT	GGG	AAG	TCT	AGC	TTT	TTC	AAC	TCA	GTG	AGG	TCT	GTT	TTC
30			6/	10		4	EA			660			٤٠	
	CAA	CCC	CAT	CTA	ACG	CAT	CYC	CCT	dedic.	CTG	ccc	A CT	D A R A	7U 2C2
													~~~	
35		•	680			690			7	00		•	710	
	ACT	GGG	ATA	TCT	GAG	AAG	TAT	AGG	ACA	TAC	TCT	ATT	AGA	GAC
		720			7	20			740			750		
40	GGG	120	GAT	GGC	7.	TAC	CTG	CCA	- Annah A A	y dada	CTC	75U	GAC	TC N
40						•••					C10	igi	GAC	TCA
	7	60		•	770			780			7	90		
	CTG	GGG	CTG	AGT	GAG	AAA	GAA	GGC	GGC	CTG	TGC	ATG	GAT	GAC
45														
							_							_
	800							20	•		830			840
	ATA	TCC	TAC	ATC	TTG	AAC	GGT	AAC	ATT	CGT	GAT	AGA	TAC	CAG
50														
₩			A	50		1	860			870)		Q	80
	TTT	AAT	CCC											

	890 ATT GAT TCC CCA		910 GAC AGA ATT	920 CAT TGT GTG GCA
8	930 TTT GTA TTT GAT		950 ATT GAA TAC	
10	970 9 ATG ATA GTA AAG	80 ATC AAA AGA		1000 GAG TTG GTA AAC
15	1010 1020 GCT GGT GTG GTA		TTG CTC ACT	
20	1060 ATG GAT CTG ATT	1070 ACA AAA GGT	1080 GAC CTT ATA	GAA ATA GAG AGA
25	1100 TGT GTG CCT GTG	1110 AGG TCC AAG	CTA GAG GAA	1130 GTC CAA AGA AAA
20	1140 CTT GGA TTT GCT	1150 CTT TCT GAC	1160 ATC TCG GTG	1170 GTT AGC AAT TAT
30	1180 1: TCC TCT GAG TGG	190 GAG CTG GAC	1200 CCT GTA AAG	1210 GAT GTT CTA ATT
35	1220 1230 CTT TCT GCT CTG	AGA CGA ATG	40 CTA TGG GC	250 1260 F GCA GAT GAC TTC
40	1270 TTA GAG GAT TTG	1280 CCT TTT GAG	1290 CAA ATA GG	1300 G AAT CTA AGG GAG
	1310 GAA ATT ATC AAC	1320 TGT GCA CAL	1330 A GGA AAA AA	A 3' - 5'

wherein the sign "-" represents a base complementary to the base shown just above each sign.

45

^{3.} An expression vector in which a DNA fragment containing a base sequence according to claim 1 or 2 and coding for an antigen specifically occurring in a host affected with non-A non-B hepatitis is introduced into a cloning site present downstream from a promoter of said vector.

^{4.} The expression vector in accordance with claim 3, in which the promoter is controllable by a regulatory factor.

^{5.} The expression vector in accordance with claim 3, in which the promoter operates in a microorganism.

^{6.} The expression vector in accordance with claim 3, in which the promoter operates in an eukaryote.

^{7.} A transformant obtained by transforming a host with an expression vector in which a DNA fragment containing a base sequence according to claim 1 or 2 and coding for an antigen specifically occurring in a host affected with non-A non-B hepatitis is introduced into a cloning site present downstream from a promoter of said vector.

8. The transformant in accordance to claim 7, in which the host is Escherichia coli or Bacillus subtills.

9. A process for producing an antigen occurring specifically in a host affected with non-A non-B hepatitis, comprising introducing a DNA fragment containing a base sequence according to claim 1 or 2 and coding for said specifically occurring antigen into a cloning site present downstream from a promoter of a vector for expression, introducing the expression vector containg said DNA fragment into a host, culturing said transformed, and collecting the produced and accumulated antigen.

10. Process for the <u>in vitro</u> diagnosis of non-A non-B hepatitis which comprises contacting a liver sample and/or a serum sample taken from a patient suspected of being infected with a non-A non-B hepatitis, with the protein whose sequence appears in claim 1, or a part thereof, for a time and under conditions sufficient to allow for the production of a complex between said protein or a part thereof with the antibodies contained in the patient sample and detecting the presence of the immunological complex, particularly when the patient is suffering from non-A non-B hepatitis.

15 Revendications

1. Fragment d'ADN qui contient une séquence de bases codant pour une protéine antigénique rencontrée spécifiquement chez un hôte atteint de l'hépatite non-A non-B, ladite protéine comprenant la totalité ou une partie de la séquence d'acides aminés représentée par la formule :

20

25

30

^=

40

45

50

	14 - A	• • •	17.3		4	λrg	T		•	10	m1 -	~ 9	7	71.	1
5	Met	YIE	AST	THE	TAP	Arg	ren	THE	ITD	Leu	H72	GTA	rys	7.70	red
•	_ •				20	41	•		_	_	_	•		•	30
	Gln	Asa	H1s	Phe	GIÅ	Gly	rys	Arg	Lau	Ser	Leu	Leu	TYE	rya	CTA
							_			40					•
10	Ser	Val	His	GIY	Phe	His	AAD	GTA	Val	Leu	Lau	Yab	Arg	CAR	CAR
		_		_	50	_	_			_•	_		_		60
	λsn	Gla	Gly	Pro	Thr	Leu	Thr	Val	Ile	Tyr	Ser	Glu	Asp	His	Ile
15										70					
	Ile	Gly	Ala	Ili	lla	Glu	Glu	Gly	TYE	Glo	Glu	Arg	Lys	IYE	Ala
					80										90
20	Ser	Ila	Ile	Leu	Phe	Ala	Leu	Gln	Glu	Thi	Lys	Ila	Ser	Glu	IIP
20										100	•				
	Lys	Lev	Gly	Leu	IAT	The	Pro	Glu	The	Lov	Phe	Cyr	Cys	Asp	Val
					110										120
25	Ala	Ly	Tyr) Ann	Ser	Pro	Thi	Asn	Phe	Gl	Ila) Asj	, Cly) Arg	Asn
										130	.				
	λro	LV	s Va	l Il	a Met	. Ast	Let	Ly:	Th			ı Ası	a Lev	ı Gly	Leu
30	•	, –,				_									
	Ali	. Gl	n ls	n Cvi	140 R Thi) c Ile	Se:	e Il	e Gl	n As	D Ty	r Gl	u Vai	l Pho	150 Arg
				. 04											
	Cv	s 61	11 - L e	n Sa	r Lei	ı laı	a Gl	ı Ar	G Lv	16 g Il		m Gl	v Va	1 11	e Glu
35	~J.		M NO	y co.					3 -1	· •		· •	,		
	7 0		T	. Ca	170		. 54	- 11	a t.a	A-	· ·	un Med	~ 61	n Dr	180 Tyr
	76)	ų AI	g rå	# 2 u	T PG	u ve	u 04.				.g	3	- 41	w FL	m rir
40		_	_	••-			_ 77	. 1-	Ti	19			61	v D-	o Ile
	G1	y 54	I Lo	u va	T GT	n Gl	H 11	- AL	A TT		IL LA		in at	I ST	

0

210 200 (Gly Ala Gly Lys Ser Ser Phe Phe Asn Ser Val Arg Ser Val Phe 220 Gln Gly His Val Thr His Gln Ala Leu Val Gly Thr Asn Thr Thr 230 240 Gly Ile Ser Glu Lys Tyr Arg Thr Tyr Ser Ile Arg Asp Gly Lys Asp Gly Lys Tyr Leu Pro Phe Ile Leu Cys Asp Ser Leu Gly Leu 270 Ser Glu Lys Glu Gly Gly Leu Cys Met Asp Asp Ile Ser Tyr Ile 280 Leu Asn Gly Asn Ile Arg Asp Arg Tyr Gln Phe Asn Pro Het Glu 290 Ser Ile Lys Leu Asn His His Asp Tyr Ile Asp Ser Pro Ser Leu 310 25 Lys Asp Arg Ile His Cys Val Ala Phe Val Phe Asp Ala Ser Ser 320 Ile Glu Tyr Phe Ser Ser Gln Het Ile Val Lys Ile Lys Arg Ile 30 Arg Arg Glu Leu Val Asn Ala Gly Val Val His Val Ala Leu Leu 370 Glu Ile Glu Arg Cys Val Pro Val Arg Ser Lys Leu Glu Glu Val 390 Gln Arg Lys Leu Gly Phe Ala Leu Ser Asp Ile Ser Val Val Ser 400 Asn Tyr Ser Ser Glu Trp Glu Leu Asp Pro Val Lys Asp Val Leu 410 Ile Leu Ser Ala Leu Arg Arg Met Leu Trp Ala Ala Asp Asp Pha 80 430 Leu Glu Asp Leu Pro Phe Glu Gln Ile Gly Asn Leu Arg Glu Glu Ile Ile Asn Cys Ala Gln Gly Lys Lys ***.

^{2.} Fragment d'ADN selon la revendication 1, dans lequel la séquence de bases comprend la totalité ou une partie de la séquence de bases représentée par la formule :

			1	10			20			30			4	10
	5' ATG	GCA	GTG	ACA	ACT	CGT	TTG	ACA	TGG	TIG	CAT	GAA	AAG	ATC
	3'	•••												
5			50			60				70			20	
	CTG	CAA	AAT	CAT	TTT	GGA	GGG	MG	CGG	/U	160	Color	CTC	717
		•••												
10		90			10	00		1	110			120		
	AAG	GGT	AGT	GTC	CAT	GGA	TTC	CXT	AAT	GGA	GTT	TTG	CTT	GAC
													•••	•••
	15	30		•	140		•	150			1	50		
15) AGA	TGT	TGT	AAT	CAA	GGG	CCT	ACT	CTA	AC1	GTG) }TT	TAT	AGT
				•••										
							_							
	170			180			1	90			200			210
20	GAA	GAT	CAT	ATT	ATT	GGA	GCX	TAT	GCA	GXX	GAG	GGT	TAC	CAG
20														
			2:	20			230			240			2	50
	GAA	AGA	AAG	TAT	GCT	TCC	ATC	ATC	CIT	TIT	GCA	CTT	CAA	GAG-
25			***										•••	
20			260			270			_				•	
	100		260	•		270 TGG							290	
	VCI	~~~		100			~~			CIA	TAT	ACA	CCA	GAA
20		_ ======						- -						
30		300				10						330		
	101	-			800	010	CTO	663	222	-				1.00
	NCA	CTG	111	101	TGT	GAC	ALT	GCV		TAT	AAC	100	CCY	ACT

	34	10		3	50			360			37	0		
			CAG	_	_		AGA	AAT	AGA		GTG	ATT	ATG	GAC
													•••	
5					•									
	380						40					_		420
	TTA	AAG	λCλ	ATG	GAA	AAT	CTT	GGA	CTT	GCT	CYY	AAT	TGT	ACT
			•••			•••								
				_										
10				30			140	•					46	-
	ATC	TCT	ATT	CAG	GAT	TAT	GAA	GTT	TTT	CGA	TGC	GAA	GAT	TCA
	-										•••		-	
			70			490			40				200	
15	000		70											116
	CTG	GAC	GAA	AGA .	DAA	ATA	777	GGG	GIC	ATT	GAG	CTC	AUG	AAG
		510			52	20		•	110			540		
	AGC		CTG										GGA	TCC
20									••••					
	5:	50		5	60		·	570			58	30		
	CTG	GTT	CAA	CAA	ATA								ATT	GGA
25		•••												
	590			600			6	10			520			630
	GCT	GGG	AAG	ICI	AGC	TTT	TTC	YYC	TCA	GTG	AGG	TCT	GTT	TTC
					~~~									
30							4 T A							
	<b>611</b>	000	<b>71</b>	<b>1</b> 0	100	C1#	V50	C C C C		000	-			70
	CAA	999	CAT	GIA	ACG	CAT		GCT	TIG	UTU	GGC	ACT	AAT	ACA
<b>35</b>			680			690			7	00		•	710	
	ACT		ATA											GAC
		***			•••									
									•					
40		720			7	30			740			750		
•	GGG	XXX	GAT	GGC	XXX	TAC	CTG	CCX	TTT	ATT	CTG	TGT	GAC	TCA
		•											***	
														•
45	CTG	GGG	CTG	AGT	GAG	AAA	GAA	GGC	GGC	CTG	TGC	ATG	GAT	GAC
		-	-						-		-		-	
	000			010			•	20						<b>6</b> 4 6
	800	800	<b>M1</b> 4	810		224	-	20			830			840
50	ATA	TCC	TAL	, AIC	117				, wii		Lau .	AGA	TAC	CAG
			, <del>, _</del> = 4											
			8	50			860			870	)		A	80
	TTT	' AAT												TAC
		•					-							-
RR														

		8	90			900			91	٥		9	26	
	ATT					CTG								GCA
				•••		•••								
5		930			9.4	0		•				060		
	ተተተ ተ		TTT	GAT		AGC							TCT	CAG
						•								
				_	. – •									
10													-	
	ATG	ATA	GTA	AAG	ATC		AUA	ATT	CGA	AGG	GAG	TIG	GTA	AAC
4.00	1010													
15	GCT	GGT	GTG	GTA	CAT	GTG	GCT	TTG	CIC	ACT	CAT	GTG	GAT	AGC
			10	60		10	70		1	080		•	109	90
20	ATG	GAT				XXX								
20														
		1	100			1110			11.	20		1	120	
	TGT	_										_		AAA .
25														
		1140	***	CCT		50 #C#			160			1170	117	<b>T1</b> T
	CIT	GGA	111	001	CIT	TCT		710	100	910	GII	700	W. 1	
30	_			•										
	1	180		1	190			1200			12	10		
	TCC	TCT	GAG	. TGG	GAG	CIG	GAC	CCI	GTA	AAG	GAT	GIT	CIA	ATT
												-		
35	1220			1230			12	40		1	250			1260
		TCT								_			GAC	TTC
									•		-			
			17	70		•	290			1200			1 2	00
40	<b>ጥ</b> ን	GAG		• •	CCT	_								GAG
									-		-		-	
		_	310			1320		CCI		30	31			•
45	GAA	ATT	A10		101	GCA			• ~~		. J'			

dans laquelle le signe "--" représente une base complémentaire de la base représentée juste au-dessus de chaque signe.

3. Vecteur d'expression dans lequel un fragment d'ADN contenant une séquence de bases telle que définie à la revendication 1 ou 2 et codant pour un antigène rencontré spécifiquement chez un hôte atteint de l'hépatite non-A non-B, est introduit dans un site de clonage présent en avai d'un promoteur dudit vecteur.

*50* 

. **55** 

- 4. Vecteur d'expression selon la revendication 3, dans lequel le promoteur peut être contrôlé par un facteur de régulation.
  - 5. Vecteur d'expression selon la revendication 3, dans lequel le promoteur opère dans un microorganisme.
  - 6. Vecteur d'expression selon la revendication 3, dans lequel le promoteur opère dans un eucaryote.
- 7. Transformant obtenu par transformation d'un hôte par un vecteur d'expression dans lequel un fragment d'ADN contenant une séquence de bases telle que définie à la revendication 1 ou 2 et codant pour un antigène

rencontré spécifiquement chez un hôte atteint de l'hépatite non-A non-B, est introduit dans un site de clonage présent en avail d'un promoteur dudit vecteur.

- 8. Transformant selon la revendication 7, dans lequel l'hôte est Escherichia coli ou Bacilius subtilis.
- 9. Procédé de production d'un antigène rencontré spécifiquement chez un hôte atteint de l'hépatite non-A non-B, comprenant l'introduction d'un fragment d'ADN contenant une séquence de bases telle que définie à la revendication 1 ou 2 et codant pour ledit antigène rencontré spécifiquement dans un site de clonage présent en avai d'un promoteur d'un vecteur d'expression, l'introduction du vecteur d'expression contenant ledit fragment d'ADN dans un hôte, la culture dudit transformant, et la récupération de l'antigène produit et accumulé.
- 10. Procédé pour le diagnostic in vitro de l'hépatite non-A non-B, qui comprend la mise en contact d'un échantilion de foie et/ou d'un échantilion de sérum prélevé chez un patient soupçonné être infecté par une hépatite non-A non-B, avec la protéine dont la séquence apparaît à la revendication 1, ou une partie de celle-ci, pendant un laps de temps suffisant et dans des conditions suffisantes pour permettre la production d'un complexe entre ladite protéine ou une partie de celle-ci, avec les anticorps contenus dans l'échantilion du patient, et la détection de la présence du complexe immunologique, en particulier, lorsque le patient souffre d'une hépatite non-A non-B.

# Patentansprüche

25

30

1. DNA-Fragment, das eine für ein antigenes Protein, das spezifisch in einem mit non-A non-B Hepatitis befallenen Wirt vorkommt, kodierende Basensequenz enthält, wobei das Protein die ganze oder einen Teil der durch die Formel dargestellten Aminosäuresequenz enthält:

	Met	11-	V=1	Th =	<b>**</b> -	<b>1</b>	Ť au	m	<b></b>	10	77.1 -	<b>.</b>	T	<b>71</b> -	T
5	HEL	A.C	AGT	THE	Inr	Mrg	red	THE	rrb	Leu	H12	GIR	Lys	170	rea
	Gln	Asn	His	Phe	20 Gly	Gly	Lys	λrσ	Leu	Ser	Leu	Leu	Tyr	Lvs	30 Gly
					•	•	•	- 3							•
10	Ser	Val	His	Gly	Phe	His	Asn	Gly	Val	40 Leu	Lau	λsp	yra	Cys	Сла
					50		,								60
	Asn	Gln	Gly	Pro	Thr	Leu	Thr	Val	Ila	Tyr	Ser	Glu	yzb	His	Ile
15										70					_
	Ile	Gly	Kla	Tyr	Ala	Glu	Glu	Gly	TYE	Gln	Glu	Arg	Lys	TYE	λla
	_			_	80	- •	_								, 90
20	Ser	Ile	Ila	Lau	Phe	Ala	Leu	Gln	Glu	Thr	Lys	Ile	Ser	Glu	TIP
20										100					
•	Lys	Leu	Gly	Leu	Tyr	Thr	Pro	Glu	Thr	Leu	Phe	Cys	CAz	Asp	Val
					110		•								120
25	Ala	Lys	TYE	Asn	Ser	Pro	Thr	Asn	Phe	Gla	Ile	Asp	Gly	yrg	Asn
										130	ı				
	Arg	Lys	Val	Ile	Met	Asp	Leu	Lys	Thi			i Asi	ı Lev	Gly	Leu
30					140										150
	Ala	Gli	ı Ast	Cys	140 Thr	Ile	Ser	: Ile	Gli	a Asy	Ty	: Gli	ı Val	. Phe	Arg
				_						160	•				
35	Cys	Glu	ı Asp	Ser	Leu	Asp	Glu	Arg	g Ly:			s G1;	y Val	L Ile	a Glu
•			-						_		•		•		
	Leu	Arg	Lys	s Sei	170 Leu		ı Sei	. Al	a Le	u Ar	g Th	r Ty	r Gl	ı Pr	180 Tyr
		_	•												•
40	Glv	r Sex	c Lev	ı Yal	l Gle	ı Glı	ı Ile	a Ar	g Il	19 a La		u Le	u Gl	y Pr	o Ile
	4								_					_	

	cGly	Ala	Gly	Lys	200 Ser	Ser	Phe	Phe	Asn	Ser	Val	<b>λ</b> rg	Ser	Val	210 Phe
5	(Gln	Gly	His	Val	Thr	His	Gln	Ala	Leu	220 Val	Gly	Thr	Asn	Thr	Thr
10	età'	Ile	Ser	Glu	230 Lys	Tyr	Arg	Thr	Tyr	Ser	Ile	λrg	Asp	Gly	240 Lys
	λsp	Gly	Lys	Tyr	Leu	Pro	Phe	Ile	Leu	250 Cys	Asp	Ser	Leu	Gly	Leu
15	Ser	Glu	Lys	Glu	260 Gly	Gly	Leu	Суз	Met	Asp	λsp	Ile	Ser	Tyr	270 Ile
	Leu	Asn	Gly	Asn	Ile	Arg	Asp	Arg	IYI	280 Gln	Phe	Asn	Pro	Met	Glu
20	Ser	Ile	Lys	Leu	290 Asn	His	His	Asp	Tyr	Ile	Asp	Ser	Pro	Ser	300 Leu
25	Lys	λsp	Arg	Ile	His	Cys	Val	. Ala	Phe	310 Val		Asp	Ala	Ser	Ser
	Ile	Glu	Tyr	Phe	320 Sex		- Glm	. Het	Ile	. Val	. Lys	Ile	Lys	. Arg	330 11e
30	Arg	λrg	r Glu	Leu	. Val	. Asn	Ala	. Gly	y Vál	340 L Val		. Val	L Ala	ı Lev	ı Leu
	Thr	His	. Val	. Asp	350 Sei		. Ası	Lev	ı Ile	a Thi	r Ly:	s G1;	y Asj	p Let	360 1 Ile
35	Glu	Ile	e Glu	. Arg	, Cys	. Vaj	L Pro	Val	l Ar	370 g Se:		s Lei	u Gl	u Gl	u Val
40	Gln	, <b>Ar</b> g	y Lys	Lev	380 1 Gly		e Al	ı Leı	u Se:	· = Asj	p Il	e Se	r Va	l Ya	390 1 Ser
	λsn	Ťvr	Sar	50-	· 61	. T	. Giv		<b>3</b>	400	40_9	•	•		
45					410										420
50	•									430	1				Phe
					440							Lev	ı Arg	, Gli	ı Glu
	***	774	. Wall	cys	WTG	PTD	Gly	, rå2	LY3	***	•				

^{2.} DNA-Fragment nach Anspruch 1, bei dem die Basensequenz die ganze oder einen Teil der durch die Formel dargestellten Phasensequenz enthält:

				1	.0			20			30			4	0
		ATG	GCA	GTG	λCA	ACT	CGT	TTG	ACA	TGG	TTG	CAT	GAA	λAG	ATC
5	3'														
•				50			60			7	<b>1</b> 0			80	
		CTG	CAA	AAT	CAT	TTT	GGA	GGG	AAG	CGG	CTT	AGC	CTT	CTC	TAT
						•••									
10									_						
,,,						10									
		AAG	GGT	AGT	GTC	CAT	GGA	TTC	CAT	AAT	GGA	GTT	TTG	CTT	GAC
15		1:	30		1	140			150			10	50		
70						CAA								TAT	AGT
		170			120			10	90			200			210
20						ATT									
20						ATT									
20				CAT	ATT	ATT	GGA	GCA	TAT	GCA	GAA	GAG	GGT	TAC	CAG
20		GAA	GAT	<b>CAT</b> 2:	ATT 	ATT	GGA	GCA  230	TAT	GCA	GAA  240	GAG	GGT	<b>Τλ</b> C 2	CAG  50
-		GAA	GAT	<b>CAT</b> 2:	ATT 	ATT	GGA	GCA  230	TAT	GCA	GAA  240	GAG	GGT	<b>Τλ</b> C 2	CAG
20		GAA	GAT	<b>CAT</b> 2:	ATT 	ATT	GGA	GCA  230	TAT	GCA	GAA  240	GAG	GGT	<b>Τλ</b> C 2	CAG  50
-		GAA	AGA	CAT 22 AAG	ATT  20 TAT	ATT	TCC	GCA 230 ATC	ATC	CTT	240 TTT	GAG GCA	CTT	TAC	CAG  50
-		GAA	AGA	22 AAG	20 TAT	ATT	TCC	GCA  230 ATC	ATC	GCA CTT	240 TTT	GAG GCA	CTT	TAC 2 CAA 290	CAG  50
25		GAA	AGA	22 AAG	20 TAT	ATT	TCC	GCA  230 ATC	ATC	GCA CTT	240 TTT	GAG GCA	CTT	TAC 2 CAA 290	CAG 50 GAG
-		GAA	AGA	AAG 260 ATT	20 TAT	GCT	TCC 270 TGG	GCA 230 ATC	ATC	GCA CTT	GAA 240 TTT 80 CTA	GAG GCA TAT	CTT	Z CAA 290 CCA	CAG 50 GAG
25		GAA ACT	AGA AAA 300	260 ATT	ATT 20 TAT	GCT GAA	TCC 270 TGG	GCA 230 ATC	ATC	GCA CTT 2 GGA 320	GAA 240 TTT 80 CTA	GAG GCA TAT	GGT CTT ACA	Z CAA 290 CCA	CAG 50 GAG 
25		GAA ACT	AGA AAA 300	260 ATT	ATT 20 TAT	GCT GAA	TCC 270 TGG	GCA 230 ATC	ATC	GCA CTT 2 GGA 320	GAA 240 TTT 80 CTA	GAG GCA TAT	GGT CTT ACA	Z CAA 290 CCA	CAG 50 GAG

	AAT	40 TTC	CAG	ATA	350 GAT	GGA	λGλ	360 AAT	λGλ	AAA	37 GTG	70 ATT	ATG	GAC
5	380 TTA	AAG	ACA	390 ATG	GAA	AAT	CTT	00 GGA	CTT	GCT	110 CAA	AAT	TGT	420 ACT
10	ATC	TCT	ATT	30 CAG	GAT	TAT	440 GAA	GTT	TTT	450 CGA	TGC	GAA	GAT	50 TCA
15	CTG	GAC	170 GAA	AGA	AAG	480 ATA	AAA	GGG	49 GTC	O ATT	GAG	CTC	AGG	AAG
20	AGC	510 TTA	CIG	TCT	52 GCC	TTG	λGλ	ACT	30 TAT	GAA	CCA	540 TAT	GGA	TCC
25	5! CTG	50 GTT	CAA	CAA	60 ATA	CGA	λ <b>TT</b>	570 CTG	CTG	CTG	58 GGT	CCA	ATT	GGA
30	590 GCT	GGG	AAG	600 TCT	AGC	TTT 	61 TTC	AAC	TCA	GTG	AGG	TCT	GTT	630 TTC
30	CAA	GGG	CAT	GTA	ACG	CAT	S50 CAG	GCT	TTG	660 GTG	GGC	ACT	AAT	70 ACA
35	ACT	GGG	80 ATA	TCT	GAG	690 AAG	TAT	AGG	7( ACA	OO TAC	TCT	ATT	710 AGA	GAC
40	GGG 	720 AAA	GAT	GGC	7: AAA	30 TAC	CTG	CCA	740 TTT	ATT	CTG	750 TGT	GAC	TCA
45	CTG	60 GGG	CTG	AGT	770 GAG	<b>AAA</b>	GAA	780 GGC	GGC	CTG	79 TGC	90 ATG	GAT	GAC
<i>50</i>	800 ATA		TAC	810 ATC	TTG	AAC	87 GGT	AAC	ATT	CGT	330 GAT	AGA	TAC	840 CAG
55	TTT 	AAT	85 CCC	50 ATG	GAA	TCA	ATC	AAA	TTA	870 AAT	CAT	CAT	88 GAC	BO TAC

		890			900			91	٥		9	20	
	ATT G	AT TCC											GCA
5													
•													
	9:	30		94	0		9	50			960		
	TTT G	IA TTT	GAT	GCC	AGC	TCT	ATT	GλA	TAC	TTC	TCC	TCT	CAC
10			_	_									
	970		9	80		•	990			100	0		
	ATG A	TA GTA	AAG .	ATC	AAA	AGA	ATT	CGA	λGG	GAG	TTG	GTA	AAC
15	1010	•	020			103						•	050
		ST GTG	LOZO GTA	C እ ጥ	GTG.	CCT.		~=~	10	40	cmc	C1m	.U5U
			31A				110	CIC	ACT	CAT	616	GAT	AGC
		106	50		10	70		1	080			109	0
20	ATG G	AT CTG	ATT .	ACA	AAA	GGT	GAC	CTT	ATA	GAA	λΤλ	GAG	AGA
		1100		1	110			112	20		11	130	-
25	TGT G	rg cct	GTG .	AGG	TCC	AAG	CTA	GAG	CAX	GTC	CAA	λGλ	XXX
	114	10	•	116			11	160					
		SA TTT	CCT	CTT3	0 TCT	GAC	ATC.	MCC.	~		170		
30			301				VIC.	1.0	GIG	GIT	AGC	AAT	TAT
30													
	1180	)	11	90		1	200			121	10		
	TCC TC	CT GAG	TGG	GAG	CTG	GAC	CCT	GTA	AAG	GAT	GTT	CTA	ATT
35		_											
	1220	1	1230			124	0		1:	250		•	1260
	CIT TO	CT GCT	CTG .	AGA	CGA	ATG	CTA	TGG	GCT	GCA	GAT	GAC	TTC
40		122	70		3 -	90		•			•		
	ጥጥክ ርያ	G GAT						1773	1290		~	13	30
			110				~~~	WIW		AAT.	CTA	AGG	GAG
		<b></b>			<del></del>	· · <del></del>							
		1310		1	320			13:	30				
45	GAA A1	TT ATC								3•			
										5'			

wobei "-" jeweils die zur direkt darüberstehenden Base komplementäre Base darstellt.

*50* 

- 3. Expressionsvektor, bei dem ein DNA-Fragment, das eine Basensequenz nach Anspruch 1 oder 2 enthäit und für ein Antigen kodiert, das spezifisch in einem mit non-A non-B Hepatitis befallenen Wirt auftritt, in eine stromabwärts vom Promotor des Vektors vorhandene Klonierungsstelle eingeführt wird.
- 4. Expressionsvektor nach Anspruch 3, bei dem der Promotor durch einen Regulationsfaktor kontrollierbar ist.
  - 5. Expressionsvektor nach Anspruch 3, bei dem der Promotor in einem Mikroorganismus wirksam ist.
  - 6. Expressionsvektor nach Anspruch 3, bei dem der Promotor in einem Eukaryonten wirksam ist.
- 7. Transformante, erhalten durch Transformieren eines Wirtes mit einem Expressionsvektor, bei dem ein DNA-Fragment, das eine Basensequenz nach Anspruch 1 oder 2 enthält, die für ein Antigen kodiert, das spe-

zifisch in einem mit non-A non-B Hepatitis befallenen Wirt auftritt, in eine stromabwärts vom Promotor des Vektors vorhandene Klonierungsstelle eingeführt wird.

- 8. Transformante nach Anspruch 7, bei der der Wirt Escherichia coli oder Bacilius subtilis ist.
- 9. Verfahren zur Herstellung eines Antigens, das spezifisch in einem mit non-A non-B Hepatitis befallenen Wirt vorkommt, umfassend die Schritte: Einführen eines DNA-Fragments, das eine Basensequenz nach Anspruch 1 oder 2 enthält, die für das spezifisch vorkommende Antigen kodiert, in eine stromabwärts vom Promotor eines Expressionsvektors vorhandene Klonierungsstelle, Einführen des das DNA-Fragment enthaltenden Expressionsvektors in einen Wirt, Kultivieren der Transformanten und Gewinnung des produzierten und akkumulierten Antigens.
- 10. Verfahren zur in vitro Diagnose von non-A non-B Hepatitis, umfassend die Schritte : in Kontakt bringen einer Leberprobe und/oder einer Serumprobe, die von einem Patienten mit Verdacht auf eine non-A non-B Hepatitisinfektion entnommen wurde, mit dem Protein mit einer Sequenz wie in Anspruch 1, oder einem Teil davon, für eine bestimmte Zeit und unter Bedingungen, die ausreichend sind für die Komplexbildung zwischen dem Protein, oder einem Teil davon, mit dem in der Patientenprobe enthaltenen Antikörpern, und Nachweis der Anwesenheit des immunologischen Komplexes, insbesondere wenn der Patient an non-A non-B Hepatitis leidet.

10

20

25

# Fig. la

10 20 5' ATG GCA GTG ACA ACT CGT TTG ACA TGG TTG 50 60 CAT GAA AAG ATC CTG CAA AAT CAT TTT GGA 70 80 GGG AAG CGG CTT AGC CTT CTC TAT AAG GGT 100 110 120 AGT GTC CAT GGA TTC CAT AAT GGA GTT TTG 130 140 150 CTT GAC AGA TGT TGT AAT CAA GGG CCT ACT 170 160 180 CTA ACA GTG ATT TAT AGT GAA GAT CAT ATT 190 200 ATT GGA GCA TAT GCA GAA GAG GGT TAC CAG 220 230 240 GAA AGA AAG TAT GCT TCC ATC ATC CTT TTT 250 260 GCA CTT CAA GAG ACT AAA ATT TCA GAA TGG 280 290 300 AAA CTA GGA CTA TAT ACA CCA GAA ACA CTG

# Fig. 1b

310 320 330 TTT TGT TGT GAC GTT GCA AAA TAT AAC TCC 340 350 360 CCA ACT AAT TTC CAG ATA GAT GGA AGA AAT 370 380 390 AGA AAA GTG ATT ATG GAC TTA AAG ACA ATG 400 410 420 GAA AAT CTT GGA CTT GCT CAA AAT TGT ACT 430 440 450 ATC TCT ATT CAG GAT TAT GAA GTT TTT CGA 460 470 480 TGC GAA GAT TCA CTG GAC GAA AGA AAG ATA 490 500 510 AAA GGG GTC ATT GAG CTC AGG AAG AGC TTA 520 530 540 CTG TCT GCC TTG AGA ACT TAT GAA CCA TAT 550 560 570 GGA TCC CTG GTT CAA CAA ATA CGA ATT CTG 580 590 600 CTG CTG GGT CCA ATT GGA GCT GGG AAG TCT

# Fig. lc

AGC TTT TTC AAC TCA GTG AGG TCT GTT TTC CAA GGG CAT GTA ACG CAT CAG GCT TTG GTG GGC ACT AAT ACA ACT GGG ATA TCT GAG AAG TAT AGG ACA TAC TCT ATT AGA GAC GGG AAA GAT GGC AAA TAC CTG CCA TTT ATT CTG TGT GAC TCA CTG GGG CTG AGT GAG AAA GAA GGC GGC CTG TGC ATG GAT GAC ATA TCC TAC ATC 820 830 TTG AAC GGT AAC ATT CGT GAT AGA TAC CAG TTT AAT CCC ATG GAA TCA ATC AAA TTA AAT 880 890 900 CAT CAT GAC TAC ATT GAT TCC CCA TCG CTG

# Fig. 1d

910 920 AAG GAC AGA ATT CAT TGT GTG GCA TTT GTA 950 940 TIT GAT GCC AGC TCT ATT GAA TAC TTC TCC 970 980 . TCT CAG ATG ATA GTA AAG ATC AAA AGA ATT 1000 1010 1020 CGA AGG GAG TTG GTA AAC GCT GGT GTG GTA 1030 1040 1050 CAT GTG GCT TTG CTC ACT CAT GTG GAT AGC 1070 1060 1080 ATG GAT CTG ATT ACA AAA GGT GAC CTT ATA 1090 1100 1110 GAA ATA GAG AGA TGT GTG CCT GTG AGG TCC 1120 1130 1140 AAG CTA GAG GAA GTC CAA AGA AAA CTT GGA 1150 1160 1170 TTT GCT CTT TCT GAC ATC TCG GTG GTT AGC 1180 1190 1200 AAT TAT TCC TCT GAG TGG GAG CTG GAC CCT

# Pig. le

1210 1220 1230
GTA AAG GAT GTT CTA ATT CTT TCT GCT CTG

1240 1250 1260 AGA CGA ATG CTA TGG GCT GCA GAT GAC TTC

1270 1280 1290 TTA GAG GAT TTG CCT TTT GAG CAA ATA GGG

1300 1310 1320

AAT CTA AGG GAG GAA ATT ATC AAC TGT GCA

1330 CAA GGA AAA AAA 3' 5'

Fig. 2

AAAAATTTATTTGCTTTCAGGAAAATTTTTCTGT TTTTTAAATAAACGAAAGTCCTTTTAAAAAAGACA

ATAATGTGTGGAATTGTGAGCGGATAACAATTTC TATTACACACCTTAACACTCGCCTATTGTTAAAG

600 CAA Gln

TTT TTC AAC TCA GTG AGG TCT GTT Phe Phe Asn Ser Val Arg Ser Val

570 AGC Ser

550 560 CTG GGT CCA ATT GGA GCT GGG AAG TCT Leu Gly Pro Ile Gly Ala Gly Lys Ser

TTC

# Fig. 3

300 CCA Pro	360 GAA G1u	420 TGC Cyb	480 CTG Leu	540 CTG
TCC	ATG	420 CGA TGC Arg Cys	TTA	CTG
AAC Asn	ACA	TTT	AGC	ATT
290 TAT TYF	NSO AAG Lys	GTT GTT Val	70 AAG Lys	30 CGA Fra
AAA Lys	TTA	S A n	AGG Arg	5 ATA Ile
290 GAC GTT GCA AAA TAT AAC ASP Val Ala Lys Tyr Asn	GAC	370 380 390 400 400 410 CTT GGA CTT GCT CAA AAT TGT ACT ATC TCT ATT CAG GAT TAT GAA GTT Leu Gly Leu Ala Gln Asn Cys Thr Ile Ser Ile Gln Asp Tyr Glu Val	CTC	530 CAA ATA CGA ATT CTG CTG Gln Ile Arg Ile Leu Leu
80 GTT Val	NO ATG Met	OCAT ASP	GAG Glu	O CAA Gln
CAC Asp	ATT Ile	CAG Gla	46 ATT Ile	52 GTT Val
TGT Cys	GTG. Val	ATT	GTC Val	CrG
TGT Cys	AAA Lys	TCT	666 617	TCC
270 TTT Phe	330 AGA Arg	390 ATC Ile	450 AAA Lys	510 GGA G1y
CTG Leu	AAT Asn	ACT	ATA Ile	TAT
ACA	AGA	TGT Cys	AAG Lys	CCA
260 GAA G1u	320 GGA G1y	AAT ABD	AGA AGA Arg	GAA GAA Glu
S C C	GAT Asp	្ត្រ <b>ខ្</b> ពុខ	ega Glu	S TAT TYF
ACA	ATA Ile	GCT	GAC	ACT
SO TAT TYF	CS GIn GIn	CTT	CTG	0 AGA Arg
CTA Leu	TTC Phe	37 668 614	TCA Ser	49 TTG
GGA CTA TAT ACA CCA GAA ACA CTG TTT TGT TGT G1y Leu Tyr Thr Pro Glu Thr Leu Phe Cys Cys	310 350 330 340 350 350 AAT AGA AAT AGA AAA GTG ATT ATG GAC TTA AAG ACA ATG GAA Asn Phe Gln ile Asp Gly Arg Asn Arg Lys Val ile Met Asp Leu Lys Thr Met Glu	CIT	GAT TCA CTG GAC GAA AGA AAG ATA AAA GGG GTC ATT GAG CTC AGG AAG AGC TTA A8P Ser Leu Asp Glu Arg Lys Ile Lys Gly Val Ile Glu Leu Arg Lys Ser Leu	GCC TTG AGA ACT TAT GAA CCA TAT GGA TCC CTG GTT CAA Ala Leu Arg Thr Tyr Glu Pro Tyr Gly Ser Leu Val Gln
CTA	ACT	AAT	GAA	TCT

CGA

CAG ATG ATA GTA AAG ATC AAA AGA ATT Gln Met Ile Val Lys Ile Lys Arg Ile

930 TCT Ser

**JCC** 

920

910

GAT GCC AGC TCT ATT GAA TAC TTC ASP Ala Ser Ser Ile Glu Tyr Phe

Ser

1010 1020 CAT GTG GAT AGC ATG His Val Asp Ser Met

TTG CTC ACT Leu Leu Thr

GCT

GTG

GTA Val

AGG GAG TTG GTA AAC GCT GGT GTG Arg Glu Leu Val Aen Ala Gly Val

980

970

990 CAT His

1010

1000

Asp

960

3p Fig.

i

660 TAT TYE	720 GAC ASP	780 TTG Leu	840 CAT His	900 TTT Phe
AAG Lys	TGT Cyb	ATC Ile	AAT Asn	GTA Val
GAG AAG	720 CTG TGT GAC Leu Cys Asp	TAC	TTA AAT Leu Asn	TTT GTA
	10 ATT Ile	70 TCC Ser	30 AAA Lys	90 GCA Ala
ATA Ile	TTT Phe	ATA Ile	ATC Ile	GTG Val
666 61y	CCA	GAC	830 TCA ATC AAA 1 Ser Ile Lys I	TGT Cys
10 ACT Thr	CTG Lau	CAT ASP	O GAA Glu	O CAT His
ACA Thr	70 TAC TYE	76 ATG Met	82 ATG Met	88 ATT Ile
AAT Asn	AAA	TGC	CCC	AGA
ACT	66C 61y	CTG	AAT Asn	GAC
630 GGC GLY	690 Gat Abp	750 GGC GLY	810 TTT Phe	870 AAG Lys
GTG Val	AAA Lys	GGC	CAG Gln	CTG Leu
TTG	666 61y	GAA	TAC	TCG Ser
520 GCT Ala	670 680 690 710 TAC TCT ATT AGA GAC GGG AAA GAT GGC AAA TAC CTG CCA TTT ATT TYr Ser Ile Arg Asp Gly Lys Asp Gly Lys Tyr Leu Pro Phe Ile	740 AAA Lys	AGA Arg	60 CCA Pro
CAG Gln	AGA Arg	GAG Glu	GAT Asp	a TCC Ser
CAT	ATT Ile	AGT	CGT	gat Asp
LO ACG Thr	70 TCT Ser	CTG Leu	ATT Ile	0 ATT Ile
6 GTA Val	TAC Tyr	73 666 614	79 AAC Asn	85 Tac
GGG CAT GTA ACG CAT CAG GCT TTG GTG GGC ACT AAT ACA ACT GGG ATA TCT GTY His Val Thr His Gln Ala Leu Val Gly Thr Asn Thr Thr Gly Ile Ser	AGG ACA Arg Thr	730 740 750 760 760 770 780 CTG GGG CTG AAA GAA GGC GGC CTG TGC ATG GAT GAC ATA TCC TAC ATC TTG Leu Gly Leu Cys Met Asp Asp Ile Ser Tyr Ile Leu	GGT AAC ATT CGT GAT AGA TAC CAG TTT AAT CCC ATG GAA Gly Asn ile Arg Asp Arg Tyr Gln Phe Asn Pro Met Glu	GAC TAC ATT GAT TCC CCA TCG CTG AAG GAC AGA ATT CAT TGT GTG GCA Asp Tyr ile Asp Ser Pro Ser Leu Lys Asp Arg ile His Cys Val Ala
666 61y	AGG	TCA	AAC	CAT
	-			

1200 CTG AGA Leu Arg

1190 CTT TCT Leu Ser

GTT CTA ATT Val Leu Ile

Asp

Lys

GAT

AAG

1170

CCT GTA

1150 TCT GAG TGG GAG CTG GAC Ser Glu Trp Glu Leu Asp

TCC

TAT

Tyr

1180

GCT

1260 ATA GGG AAT Ile Gly Asn

GAG CAA Glu Gln

TTG CCT TTT Leu Pro Phe

GAT Asp

GAG Glu

1230 TTC TTA Phe Leu

CGA ATG CTA TGG GCT GCA GAT GAC Arg Met Leu Trp Ala Ala Asp Asp

1250

1240

Fig. 3c

1080 AAG Lys	1140 AAT Asn
TCC Ser	AGC Ser
AGG	GTT Val
1070 CCT GTG Pro Val	1130 TCG GTG Ser Val
CCT	
GTG	ATC Ile
50 TGT Cys	20 GAC ABP
AGA TG Arg Cy	1120 TCT GA Ser As
GAG Glu	Crr
ATA Ile	GCT
1050 GAA Glu	1110 TTT Phe
ATA Ile	GGA G1y
CTT	Cir
1040 GGT GAC Gly Asp	AGA AAA Arg Lys
GGT	AGA Arg
AAA	CAA G1n
1030 ATT ACA Ile Thr	1090 GAA GTC Glu Val
10 ATT Ile	10 68 61 61 61
region .	GAG Glu
GAT	CTA

1270 1300 CTA AGG GAG GAA ATT ATC AAC TGT GCA CAA GGA AAA AAA TAG Leu Arg Glu Glu Ile Ile Asn Cys Ala Gln Gly Lys Lys ***

AAGATAAAAG GGGTCATTGA GCTCAGGAAG TTCTATTTTC CCCAGTAACT CGAGTCCTTC

GGACGAAAGA Z

520 AAGATTCACT TTCTAAGTGA

CTATTCAGGA TTATGAAGTT TTTCGATGCG
GATAAGTCCT AATACTTCAA AAAGCTACGC

530

# Fig. 4a

80	160	240	320	400	480
TCGTTTGACA	TCCATGGATT	CATATTATTG	TAAAATTTCA	CTAATTTCCA	TGTACTATCT
AGCAAACTGT	AGGTACCTAA	GTATAATAAC	ATTTTAAAGT	GATTAAAGGT	ACATGATAGA
70	150	230	310	390	470
CAGTGACAAC	AAGGGTAGTG	TAGTGAAGAT	TTCAAGAGAC	AACTCCCCAA	TGCTCAAAAT
GTCACTGTTG	TTCCCATCAC	ATCACTTCTA	AAGTTCTCTG	TTGAGGGGTT	ACGAGTTTTA
60	140	220	300	380	460
AGAAGTATGG	CCTTCTCTAT	CAGTGATTTA	CTTTTGCAC	TGCAAAATAT	ATCTTGGACT
TCTTCATACC	GGAAGAGATA	GTCACTAAAT	GAAAAACGTG	ACGTTTTATA	TAGAACCTGA
50	130	210	290	370	450
CAACAGATCA	AGCGGCTTAG	CCTACTCTAA	TTCCATCATC	GTTGTGACGT	ACAATGGAAA
GTTGTCTAGT	TCGCCGAATC	GGATGAGATT	AAGGTAGTAG	CAACACTGCA	TGTTACCTTT
40 ACAGACAGTA TGTCTGTCAT	120 TTTGGAGGGA	200 TAATCAAGGG ATTAGTTCCC	280 GAAAGTATGC CTTTCATACG	360 ACACTGTTTT TGTGACAAAA	440 GGACTTAAAG CCTGAATTTC
30	110	190	270	350	430
AGCTCATACT	GCAAAATCAT	ACAGATGTTG	TACCAGGAAA	TACACCAGAA	AAGTGATTAT
TCGAGTATGA	CGTTTTAGTA	TGTCTACAAC	ATGGTCCTTT	ATGTGGTCTT	TTCACTAATA
20	100	180	260	340	420
CCTCAGCTCT	AAAAGATCCT	GTTTGCTTG	AGAAGAGGGT	TAGGACTATA	AGAAATAGAA
GGAGTCGAGA	TTTTCTAGGA	CAAAACGAAC	TCTTCTCCCA	ATCCTGATAT	TCTTTATCTT
10	90	170	250	330	410
GGGGGGCTAC	TGGTTGCATG	CCATAATGGA	GAGCATATGC	GAATGGAAAC	GATAGATGGA
CCCCCCGATG	ACCAACGTAC	GGTATTACCT	CTCGTATACG	CTTACCTTTG	CTATCTACCT
. m					

880

870

860

850

840

830

820

810

CCTGTATGAG ATAATCTCTG CCCTTTCTAC CGTTTATGGA CGGTAAATAA

ACGGTAACAT

TACATCTTGA

TGACATATCC

TGTGCATGGA

GAAGGCGGCC

CTGTGTGACT CACTGGGGCT GAGTGAGAA GACACACTGA GTGACCCCGA CTCACTCTT

GATTATGTTG ACCCTATAGA CTCTTCATAT

TGCCATTGTA

640 TGGGTCCAAT ACCCAGGTTA	720 TTGGTGGGCA AACCACCCGT	800 GCCATTTATT
630 ATTCTGCTGC TAAGACGACG	AGTGAGGTCT GTTTTCCAAG GGCATGTAAC GCATCAGGCT TCACTCCAGA CAAAAGGTTC CCGTACATTG CGTAGTCCGA	790 GCAAATACCT
620 ACAAATACGA TGTTTATGCT	700 GGCATGTAAC CCGTACATTG	780 GGGAAAGATG
CCATATGGAT CCCTGGTTCA ACAAATACGA ATTCTGCTGC GGTATACCTA GGGACCAAGT TGTTTATGCT TAAGACGACG	AGTGAGGTCT GTTTTCCAAG GGCATGTAAC GCATCAGGCT TCACTCCAGA CAAAAGGTTC CCGTACATTG CGTAGTCCGA	770 TATTAGAGAC
600 CCATATGGAT GGTATACCTA	680 AGTGAGGTCT TCACTCCAGA	760 GGACATACTC
590 AACTTATGAA TTGAATACTT	670 TTTTCAACTC	750 GAGAAGTATA
580 CTGCCTTGAG GACGGAACTC	660 AAGTCTAGCT TTCAGATCGA	740 750 760 770 780 790 TGGGATATCT GAGAAGTATA GGACATACTC TATTAGAGAC GGGAAAGATG GCAAATACCT
570 AGCTTACTGT TCGAATGACA	650 TGGAGCTGGG	730 CTAATACAAC

4

Fig.

	A H	9
ATGACTACAT	1020 TTCTCCTCTC AAGAGGAGAG	1100 CACTCATGTG
TTAAATCATC ATGACTACAT AATTTAGTAG TACTGATGTA	1010 TATTGAATAC ATAACTTATG	1090 TGGCTTTGCT
TACCAGTTTA ATCCCATGGA ATCAATCAAA ATGGTCAAAT TAGGTACCT TAGTTAGTTT	TTTGTATTTG ATGCCAGCTC TATTGAATAC	1080 GTGGTACATG
ATCCCATGGA TAGGGTACCT	990 TTTGTATTTG	1070 AAACGCTGGT
TACCAGTTTA	980 TTGTGTGGCA AACACACCGT	1050 1060 1070 1080 1090 1100 AGANTTCGAA GGGAGTTGGT AAACGCTGGT GTGGTACATG TGGCTTTGCT CACTCATGTG
TCGTGATAGA AGCACTATCT	970 ACAGAATTCA TGTCTTAAGT	1050 AGAATTCGAA

ATCTGATTAC TAGACTAATG

GATAGCATGG CTATCGTACC

GTGAGTACAC

TGGCTTTGCT

GTGGTACATG

AAACGCTGGT TTTGCGACCA

CCCTCAACCA

TCTTAAGCTT

1120

1110

1040

AAAGATCAAA TITCIAGITI

1030 AGATGATAGT TCTACTATCA

960

950

TCGCTGAAGG

TGATTCCCCA

2 3

AAAAAAA AAAAAA TITITITIT ITITITI

ATAAATTTCT TATAAAAAA TATTTAAAGA ATATTTTTT

CTTGGATTTA TGTTCTGTAT CTGTGAAAA GAACCTAAAT ACAAGACATA GACCTTTTTT

	•	ŧ	٢
	•		•
			•
		ř	
		`	
1	ï	í	
		1	
	•		
	•		
	•		
	•		
	•		
	•		
	•		
	•		

1200	1280	1360	1440	1520	1600
CTTGGATTTG	AATTCTTTCT	TAAGGGAGGA	AGATTAAAAT	GATGAAGAAA	AATAATTTTT
GAACCTAAAC	TTAAGAAAGA	ATTCCCTCCT	TCTAATTTTA	CTACTTCTTT	TTATTAAAAA
1190	1270	1350	1430	1510	1590
CCAAAGAAAA	AGGATGTTCT	ATAGGGAATC	ACATCACAGA	TAATGTCTAG	GAAAAATAAT
GGTTTCTTTT	TCCTACAAGA	TATCCCTTAG	TGTAGTGTCT	ATTACAGATC	CTTTTTATTA
1180	1260	1340	1420	1500	1580
TAGAGGAAGT	GACCCTGTAA	TTTTGAGCAA	AAATTTCCTC	TGTGTTTTAT	TCATAATTGT
ATCTCCTTCA	CTGGGACATT	AAAACTCGTT	TTTAAAGGAG	ACACAAAATA	AGTATTAACA
1170	1250	1330	1410	1490	1570
AGGTCCAAGC	GTGGGAGCTG	AGGATTTGCC	AGGTTCACGT	Accaaaggga	CATGATTTAG
TCCAGGTTCG	CACCCTCGAC	TCCTAAACGG	TCCAAGTGCA	Tggtttccct	GTACTAAATC
1160	1240	1320	1400	1480	1560
TGTGCCTGTG	ATTCCTCTGA	GACTTCTTAG	GATATGTGAA	AGTAACTAAG	CTAGAAATAA
ACACGGACAC	TAAGGAGACT	CTGAAGAATC	CTATACACTT	TCATTGATTC	GATCTTTATT
1150	1230	1310	1390	1470	1550
TAGAGAGATG	GTTAGCAATT	GGCTGCAGAT	GAAAAAATA	ACCAAAGAGA	TTGTAAATAA
ATCTCTCTAC	CAATCGTTAA	CCGACGTCTA	CTTTTTTAT	TGGTTTCTCT	AACATTTATT
1140	1220	1300	1380	1460	1540
CTTATAGAAA	CATCTCGGTG	GAATGCTATG	TGTGCACAAG	GAAAACACAG	ATTGTAGTAC
GAATATCTTT	GTAGAGCCAC	CTTACGATAC	ACACGTGTTC	CTTTTGTGTC	TAACATCATG
1130	1210	1290	1370	1450	1530
AAAAGGTGAC	CTCTTTCTGA	GCTCTGAGAC	AATTATCAAC	TCAGAAAGGA	TGCATAGAAC
TTTTCCACTG	GAGAAAGACT	CGAGACTCTG	TTAATAGTTG	AGTCTTTCCT	ACGTATCTTG

Fig.5

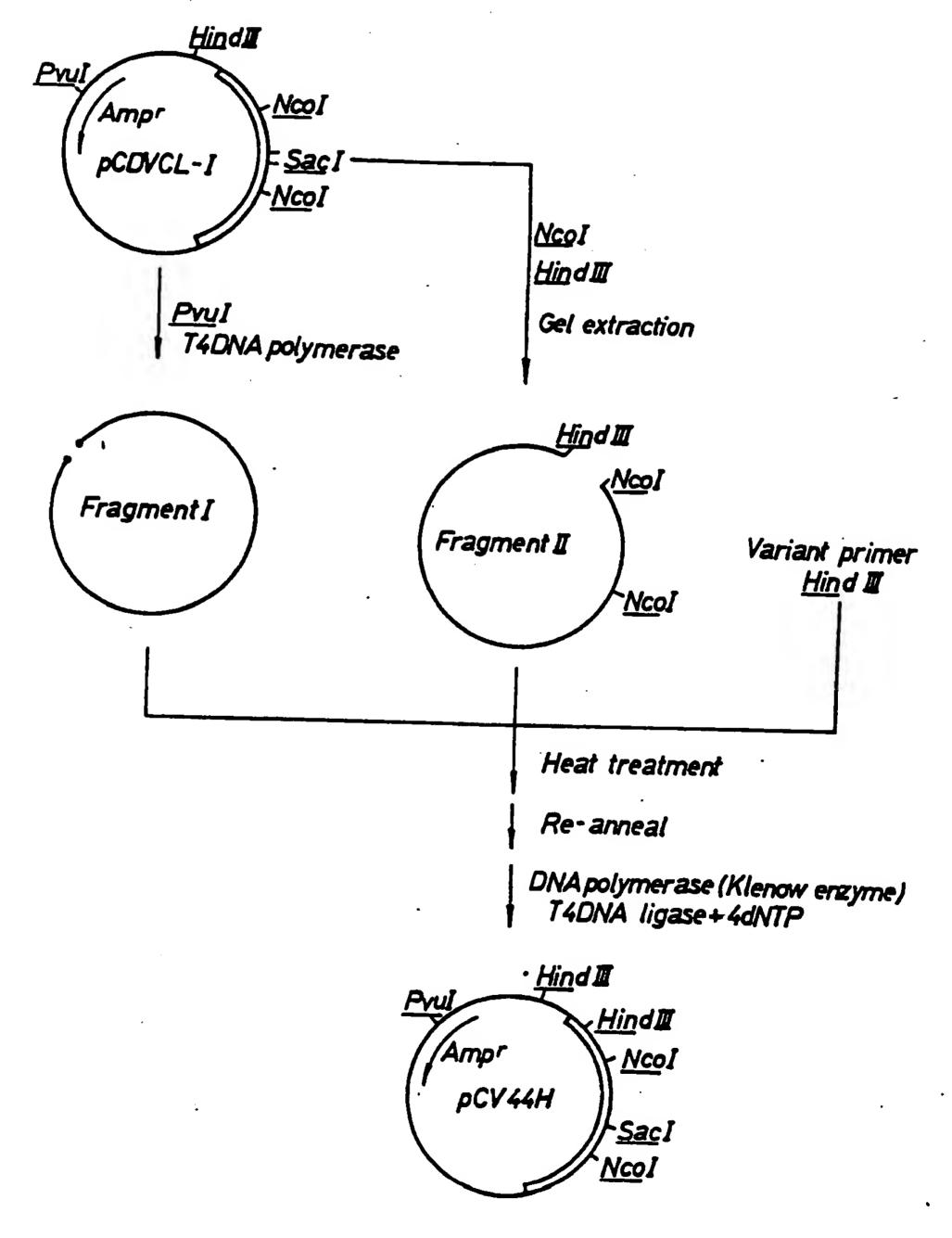


Fig.6

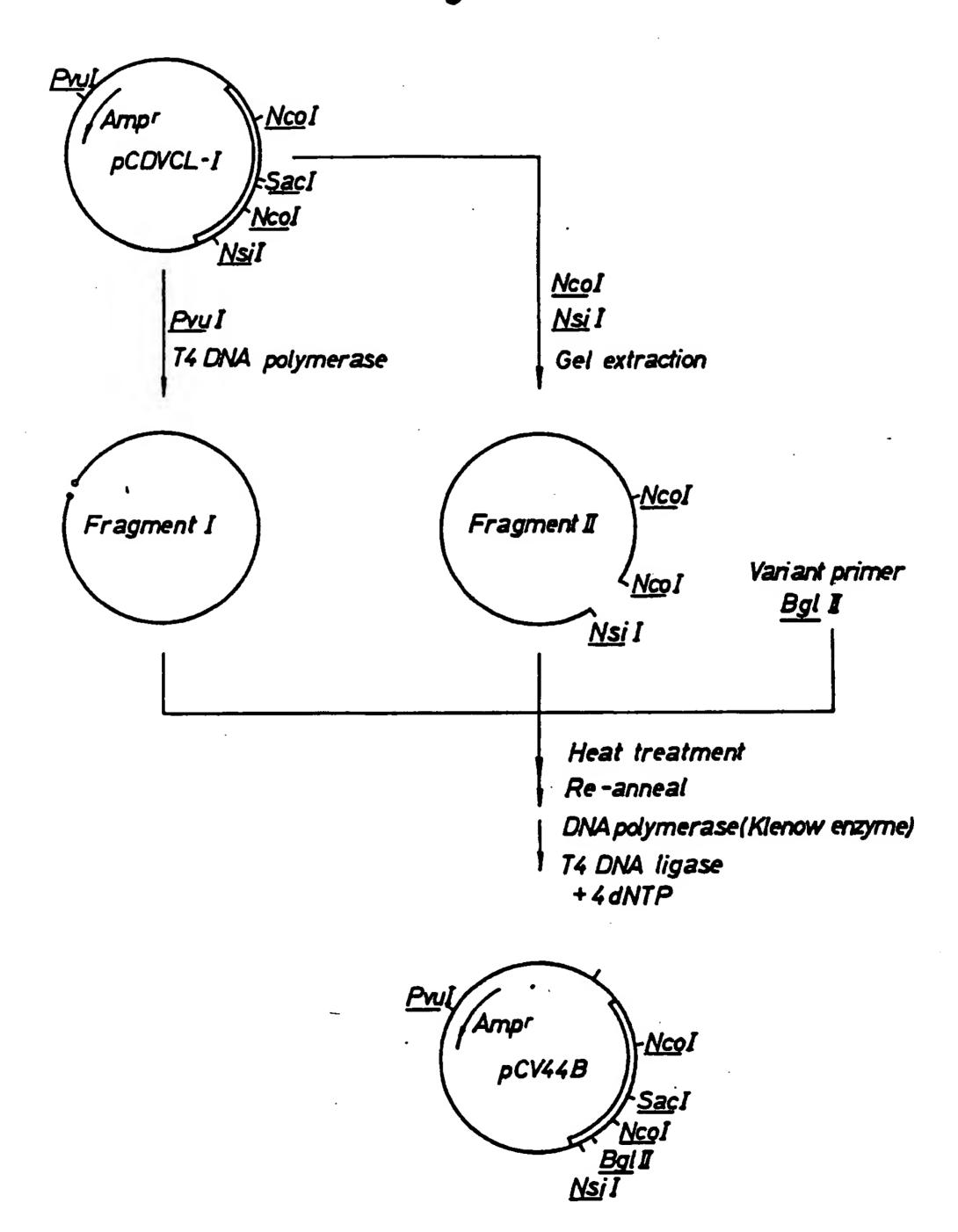


Fig.7

